Urban-Biased Growth: A Macroeconomic Analysis*

Fabian Eckert† Sharat Ganapati‡ Conor Walsh§

First Version: March 2020
This Version: April 2024

Abstract

After 1980, larger US cities experienced substantially faster wage growth than smaller ones. We show that this urban bias mainly reflected wage growth at large Business Services firms. These firms set themselves apart through high per-worker spending on information technology capital and their disproportionate presence in big cities. We introduce a spatial model of investment-specific technical change that can rationalize these patterns. Using the model as an accounting framework, the observed decline in the investment price of information technology capital explains most urban-biased growth by raising the profits of large Business Services firms.

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 Milena Almagro, Pol Antràs, David Autor, Costas Arkolakis, Adrien Bilal, Gideon Bornstein, Laura Castillo-Martinez, Jonathan Dingel, Pierre-Olivier Gourinchas, Gordon Hanson, J. Bradford Jensen, Tom Kemeny, Paolo Martellini, Christian Moser, Dávid Nagy, Michael Peters, Esteban Rossi-Hansberg, Steve Redding, and Daniel Sturm 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.

†University of California, San Diego; fpe@ucsd.edu
‡Georgetown University; sharat.ganapati@georgetown.edu
§Columbia University; caw2226@columbia.edu

INTRODUCTION

Since 1980, US wage growth has been faster in cities with higher population density. The left panel of Figure 1 shows average wages across US commuting zones grouped into deciles of increasing population density. In 1980, the average worker in the top decile, which consists of New York and Chicago, earned 32% more than the average worker in the bottom decile. By 2015, the gap had risen to 71%.

Urban-biased growth is related to many economic and societal challenges the US has faced in recent decades. It has occurred alongside skyrocketing house prices in urban centers (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, the origins of urban-biased growth remain largely unexplained.

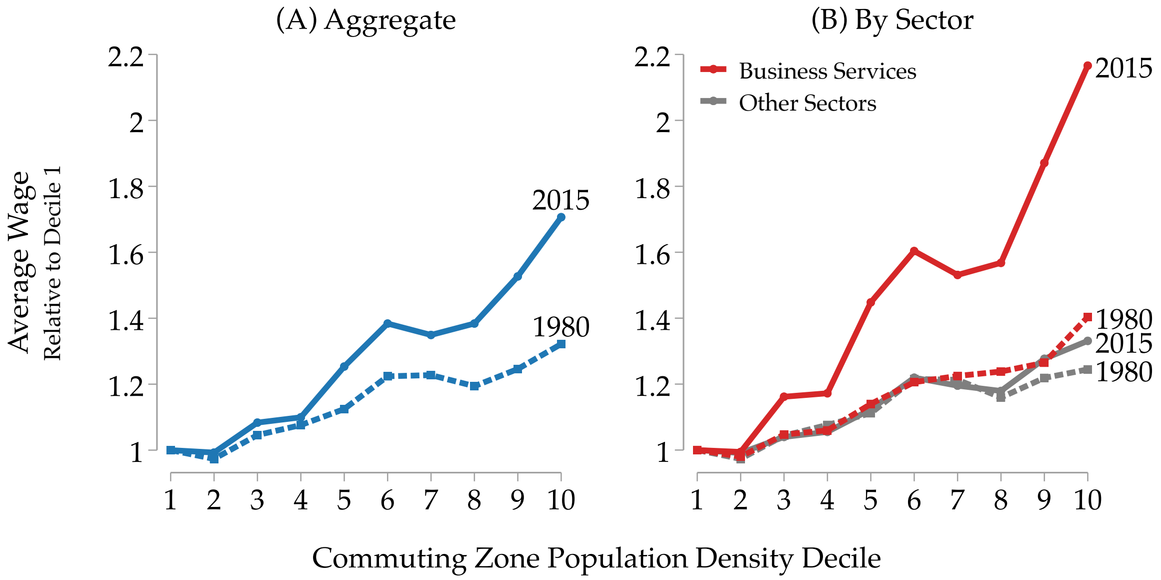
This paper uses new data and economic theory to provide an explanation for urban-biased growth. We empirically document that urban-biased growth has been driven almost entirely by large establishments in the Business Services sector (NAICS-5) that invested heavily in information technology (IT). By hosting these establishments, high-density cities have benefited more than others from the substantial decline in the price of IT capital. We then integrate investment-specific technical change and a firm-size-capital complementarity into a spatial growth model that can flexibly account for other sources of growth. We use the model for a growth accounting exercise, and find the observed decline in IT capital prices alone explains most urban-biased growth since 1980.

We begin by showing that Business Services have been responsible for virtually all urban-biased growth since 1980. The right panel of Figure 1 shows average wages across commuting zones for the Business Services sector and the rest of the economy. In 1980, Business Services workers in cities with the highest population densities earned, on average, 40% more than workers in cities with the lowest population densities. By 2015, they made 117% more. Meanwhile, the relationship between wages and population density has changed little in other sectors.

Using microdata on the universe of US establishments, we show that more than two-thirds of the urban-biased growth within the Business Services sector is due to large establishments with more than 100 employees. The outsized role of these establishments primarily reflects that wage growth was dramatically faster at large establishments in big cities than elsewhere, but also, to a lesser extent, that large establishments account for a disproportionate share of big-city Business Services employment.

Lastly, we show that Business Services establishments in high-density cities were among the largest investors in IT capital. This finding reflects a combination of two empirical regularities in the Business Services sector that are important for our theory. First, the

FIGURE 1: THE US WAGE-DENSITY GRADIENT IN 1980 AND 2015



Notes: This figure shows average annual wages across commuting zones (Tolbert and Sizer, 1996) sorted into deciles of increasing population density. Each decile accounts for one-tenth of the US population in 1980. The average commuting zone in decile 1 has a population density of 10 $people/mi^2$ and in decile 10 of 2300 $people/mi^2$. The underlying data come from the US Census Bureau's Longitudinal Business Database and cover all US private, non-farm employer establishments. We compute average wages as average payroll per worker by aggregating establishment payroll numbers and employment counts across all establishments in a commuting zone and sector. We show all wages relative to wages in decile 1.

bigger the city, the larger the average Business Services establishment. Second, the larger a firm's total employment, the higher its per-worker expenditure on IT capital. Together, these patterns point to the dramatic price decline for IT capital after 1980 as a potential driver of urban-biased growth.

We additionally show the wages of college and non-college-educated workers in Business Services have experienced urban-biased wage growth, while the wages of college- and non-college-educated workers outside the Business Services sector have not grown faster in higher-density locations.[1] These findings underline that urban-biased growth is a sectoral phenomenon not specific to an educational group, in contrast with the focus of recent literature.

We then present a dynamic quantitative spatial model of investment-specific technical change. The model shows which elasticities determine whether an aggregate decline in the price of IT capital leads to unbalanced growth across space, and under what conditions it benefits high-density locations in particular. Quantitatively, the framework enables us to isolate the contribution of the *observed* decline in IT capital prices to urban-biased growth, while flexibly accounting for other sources of spatial wage and

[1] Most college-educated workers in the US economy work outside the Business Services sector. In 2015, only 28% of all workers with a college degree worked in the Business Services sector.

employment changes.

Under general conditions, we first show an intuitive sufficient statistic governs the general equilibrium exposure of local wages to a fall in the IT investment price. The exposure statistic is the ratio of total payments to capital relative to labor across all firms within a location-sector pair. All else equal, the higher its payments to capital, the more the location-sector benefits from declines in the price of IT capital through input-cost savings. On the other hand, the lower the payments to labor, the more local wages have to increase for local firms' labor costs to increase sufficiently to make further capital investments unprofitable.

The observed decline in IT capital prices then leads to urban-biased wage growth if the exposure statistic increases with population density in the cross-section of commuting zones. In equilibrium, the spatial variation in the statistic depends on two competing channels. First, higher wages in big cities reduce exposure through a price effect and increase it through a substitution effect. All else equal, if capital and labor are complements, the price effect dominates, lowering big cities' capital price exposure. This *neoclassical channel* alone would generate rural-biased wage growth in the face of aggregate investment-specific technical change.

Our model introduces a new *scale channel* that runs counter to the neoclassical channel. To match the increasing capital intensity of large firms, we introduce a non-homotheticity into firms' production technologies that allows relative marginal products of factors to vary with firm scale. We show firms in higher-density locations produce at a larger scale if firms' entry costs require payments to a scarce local factor such as labor or land. Local entry costs and non-homotheticity imply that firm size increases with population density, and large firms produce more capital-intensively, increasing the capital price exposure of big cities.

Ultimately, whether the scale channel dominates the neoclassical channel is a quantitative question, and may vary across sectors. The key parameters governing the strength of the scale channel are (1) a non-homotheticity parameter in the production function that determines how capital intensity varies with firm size and (2) the importance of local factors in the entry cost that governs how firm size varies across locations. We discipline the non-homotheticity using micro-data on capital investments per worker across the firm size distribution. We discipline the composition of the entry cost using data on how firm size varies with local wages in the cross-section of cities.

Given the model's main elasticities, we infer the location-, sector-, and factor-specific productivity and amenity terms as structural residuals to account for the data on wages and employment counts across all US commuting zones between 1980 and 2015. We choose the productivity of IT capital production to match the time series of the investment price of IT capital.

Since our model can account for all the wage and employment variation in the data, we can decompose the observed urban-biased growth into changes due to our mechanism, and that due to changing "residual" productivity and amenity terms. To do so, we hold all productivity and amenity terms fixed at their 1980 levels, and then vary the investment price of IT capital as in the data. The decline in IT prices alone can account for most of the urban-biased wage growth in the data. Moreover, the urban-biased growth originates in the Business Services sector because our parameter estimates indicate that, in other sectors, the scale channel is weak relative to the neoclassical channel. As such, other firms in large cities do not benefit more from IT price declines than firms elsewhere.

Given our estimated labor-supply elasticities, both the quantity and the wage responses generated by the IT price decline resemble the data. In particular, the IT price decline alone leads to a substantial and urban-biased skill-deepening of the Business Services sector. Lastly, because IT capital and high-skill labor are more complementary than IT capital and low-skill labor, wages of the high-skilled workers in Business Services rise more than the wages of the low-skilled workers, replicating the compositional changes in the data. When we study the spatial patterns of changes in the productivity residuals, we find that while they have grown substantially, especially for more skilled workers, their growth exhibits virtually no spatial bias.

Literature Review. Our paper makes both empirical and theoretical contributions. The first empirical contribution is to document the steepening of the US wage-density gradient that we refer to as urban-biased growth. Related papers have studied wage convergence across US cities, that is, the relationship between initial wage levels and subsequent wage growth (see Berry and Glaeser, 2005; Moretti, 2012; Ganong and Shoag, 2017; Giannone, 2022). Others have studied the unbalanced growth of the *relative* wages of skilled and unskilled workers across cities (Beaudry, Doms, and Lewis, 2010; Baum-Snow and Pavan, 2013; Eckert, 2019; Rubinton, 2019; Moretti, 2013; Diamond, 2016). Finally, a growing literature studies how within-location inequality varies with city size or affects city neighborhoods (Davis and Dingel, 2020; Eeckhout, Hedtrich, and Pinheiro, 2021; Couture and Handbury, 2020; Almagro and Domínguez-Iino, 2022; Fogli, Guerrieri, Ponder, and Prato, 2023).

Our paper is the first to use establishment-level data to show directly that large technology, professional service, and financial firms are responsible for the recent spatially-biased growth of the US. Our sectoral perspective revises the view that big cities' recent success reflects broad-based wage growth biased toward more skilled workers.[2] Locating the urban-biased growth phenomenon in a single sector and establishment type

[2]For example, the average wages of medical doctors have grown in a remarkably balanced way across space in the same period. The same is true more generally for skilled workers not working in the NAICS-5 sector.

considerably narrows the set of potential drivers for urban-biased growth, allowing us to provide a concrete economic mechanism to explain it.

As a final empirical contribution, our paper shows that the Business Services sector is the most intensive user of IT capital in the economy, and provides direct cross-sectional evidence that IT capital expenditures in the sector are increasing in firm and city size. A large set of papers studies the role of a decline in the price of IT (or more general equipment) capital in generating *skill-biased* wage growth in the US economy (Krusell, Ohanian, Ríos-Rull, and Violante, 2000; Krueger, 1993; Lashkari, Bauer, and Boussard, 2024); our paper instead relates these price changes to the *urban-biased* growth in recent decades.[3]

On the theoretical side, we are the first to build investment-specific technical change into a spatial equilibrium model to study how aggregate changes in the investment price of capital affect wages and employment across locations. We add to a small number of papers that study capital investment in a spatial setting (Ravikumar, Santacreu, and Sposi, 2019; Anderson, Larch, and Yotov, 2020; Kleinman, Liu, and Redding, 2023; Bilal and Rossi-Hansberg, 2023), and more broadly technology adoption across space (Desmet and Rossi-Hansberg, 2014; Desmet, Nagy, and Rossi-Hansberg, 2018; Martellini, 2022; Nagy, 2023). Since this paper was first circulated, several subsequent papers have studied wage growth at headquarters establishments in big cities as a result of declining communication costs, and linked this to increases in aggregate inequality and efficiency (Kleinman, 2022, Jiang, 2023). We show explicitly that such establishments account only for a residual fraction of urban-biased growth because they contribute such a small share of overall Business Services employment.[4]

Technically, our paper embeds a non-homothetic CES production function (Sato, 1977) into the workhorse quantitative spatial model (Allen and Arkolakis, 2014; Redding, 2016; Redding and Rossi-Hansberg, 2017), and shows how the interaction of the non-homotheticity with spatial firm-size patterns gives rise to local exposure differences to investment-specific technical change. Comin, Lashkari, and Mestieri (2021) were the first to build a non-homothetic CES function into a structural macro model, using it as a utility aggregator in the study of structural change. More recently, Lashkari et al. (2024) and Trottner (2019) employed the aggregator as a production function. Our paper is particularly related to Lashkari et al. (2024), who provide direct evidence that IT capital exhibits a complementarity with firm size, which a non-homothetic CES production function captures well.

[3] Baum-Snow and Pavan (2013) is the only paper that studies an explicit capital-skill complementarity across locations by estimating local production functions similar to that in Krusell et al. (2000). However, they study how equipment capital price changes led to faster growth of the big-city college wage premium in manufacturing, a sector we document has little to do with the urban-biased wage-growth phenomenon.

[4] Headquarter services (NAICS Code 55) accounted for 2.4% of aggregate employment in 2015; the Business Services sector accounted for 26%.

1. Urban-Biased Growth in the Data

In this section, we document the urban-biased growth of the US economy between 1980 and 2015. We then offer several decompositions to shed light on the contributions of different sectors, firms, and worker types.

1.1 Main Data Sources

Our primary data source is the Longitudinal Business Database (LBD) drawn from the US Census Business Register, a database constructed from the administrative tax records of all private, non-farm employer establishments in the US. The LBD provides annual information on the total payroll and employment for these establishments between 1975 and 2015. Central to our analysis, an establishment is a *single physical location* where business is conducted, services are provided, or industrial operations are carried out. The LBD contains detailed information on the sector and location of each establishment.

Using its location identifier, we map each establishment to one of the 722 commuting zones (Tolbert and Sizer, 1996) covering the entirety of the continental US. We aggregate our data to 1-digit "NAICS" sectors designed to capture the principal functional differences between groups of industries.[5] We define a location-sector's "average wage" as the total payroll of all its establishments divided by their total employment. We deflate all nominal variables through time using the Bureau of Economic Analysis (BEA) Personal Consumption Expenditures Price Index (PCE).

1.2 Documenting Urban-Biased Growth

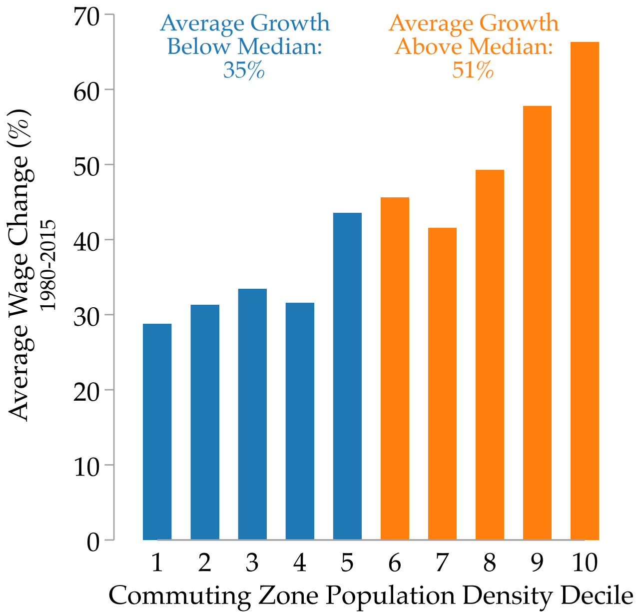
We begin by documenting the urban-biased growth of the US economy between 1980 and 2015. Based on 1980 data, we group commuting zones into deciles of increasing population density, so that each decile accounts for approximately 10% of US employment. For most of our analysis below, we compare the "low-density" commuting zones below median density with the "high-density" commuting zones above it.

Figure 2 shows the growth in average wages between 1980 and 2015 for each commuting zone decile. Average wages in the top decile of commuting zones, which includes New York and Chicago, grew twice as fast as average wages in the bottom decile. Average wages grew 51% among the above-median density commuting zones, compared to only 35% in the below-median group. In the Online Appendix, we show the urban-biased wage growth in Figure 2 represents a doubling of the wage-density gradient in the cross-section of commuting zones.

Figure 2 looks very similar when we order commuting zones by their population size

[5] NAICS stands for North American Industry Classification System. The LBD data before 1997 uses the Standard Industrial Classification (SIC) system instead. We employ the SIC-NAICS concordance from Fort and Klimek (2016) to ensure our dataset has consistent NAICS industry codes across time.

FIGURE 2: THE URBAN BIAS IN US WAGE GROWTH, 1980-2015



Notes: This figure shows wage growth between 1980 and 2015 across commuting zones (Tolbert and Sizer, 1996) sorted into deciles of increasing population density. Each decile accounts for one-tenth of the US population in 1980. The underlying data come from the US Census Bureau's Longitudinal Business Database and cover all US private, non-farm employer establishments. We compute average wages as average payroll per worker by aggregating establishment payroll numbers and employment counts across all establishments in a commuting zone and sector. All values are adjusted for inflation to 2015 dollars using the BEA PCE price index.

instead.[6] We focus on population density since population normalized by area has a more immediate economic interpretation. In addition, an extensive urban literature discusses the productive benefits and the congestion costs associated with population density (Ahlfeldt and Pietrostefani, 2019).

The urban-biased growth depicted in Figure 2 is not a unique feature of the LBD, but holds across all major US labor market datasets, including the Quarterly Census of Employment and Wages (QCEW), and the US Decennial Census. Furthermore, the European data shows similar trends, pointing to the need for an explanation not specific to the US context. Historically, we find almost no urban-biased growth between 1950 and 1980. The strong urban bias after 1980 presents a clear structural break, further narrowing the space of potential explanations.[7]

While the urban-biased wage growth in Figure 2 has received minimal attention, some related papers study other aspects of wage growth across cities. Berry and Glaeser (2005) and Giannone (2022) document "the end of wage convergence" across US cities since 1980, that is, changes in the relationship between initial wage levels and subsequent wage growth. The convergence fact and the urban-biased growth fact are separate:

[6]We present the corresponding Figure in the Online Appendix.

[7]In the Online Appendix, we present versions of Figure 2 using alternative US datasets, European data, and historical US data.

when we order commuting zones based on initial wages instead of population density, wage growth is flat across deciles, in sharp contrast with the increasing wage growth pattern in Figure 2.[8] Other related papers have studied the urban bias in the growth of the *relative* wages of skilled and unskilled workers (”skilled-wage premium”), which is separate from our focus on the level of wages (Beaudry et al., 2010; Baum-Snow and Pavan, 2013; Diamond, 2016; Eckert, 2019). A final set of papers studies how within-city wage polarization varies with city size (Davis, Mengus, and Michalski, 2020; Eeckhout et al., 2021).

1.3 Accounting for Urban-Biased Growth

In this section, we use the LBD data to shed light on the role of sectors, establishments, and IT capital in giving rise to urban-biased growth in Figure 2. We organize our findings into three facts.

Fact 1: The Business Services sector accounts for most urban-biased growth.

We first introduce a decomposition to compute the share of urban-biased growth due to a particular sector. Denote a location ℓ's average wage in sector s by $w_{\ell s}$ and a sector's share in local employment by $\mu_{\ell s}$. The difference in the growth rate of average wages between two locations ℓ and ℓ' can then be decomposed as follows:

$$(1) \qquad g_{\ell'} - g_{\ell} = \sum_s \left(\delta_{\ell' s} - \delta_{\ell s}\right) \quad \text{where} \quad \delta_{\ell s} := \frac{\mu_{\ell st+1} w_{\ell st+1} - \mu_{\ell st} w_{\ell st}}{\bar{w}_{\ell t}}$$

where $\bar{w}_{\ell t} = \sum_s \mu_{\ell st} w_{\ell st}$ and $g_{\ell} = (\bar{w}_{\ell t+1} - \bar{w}_{\ell t})/w_{\ell t}$, and a time period if indexed by t. The term $\delta_{\ell s}$ captures the (positive or negative) contribution of sector s to wage growth in location ℓ. It captures changes in wages *and* employment shares: a sector can contribute to local wage growth by generating wage increases or by growing its employment faster than other sectors.

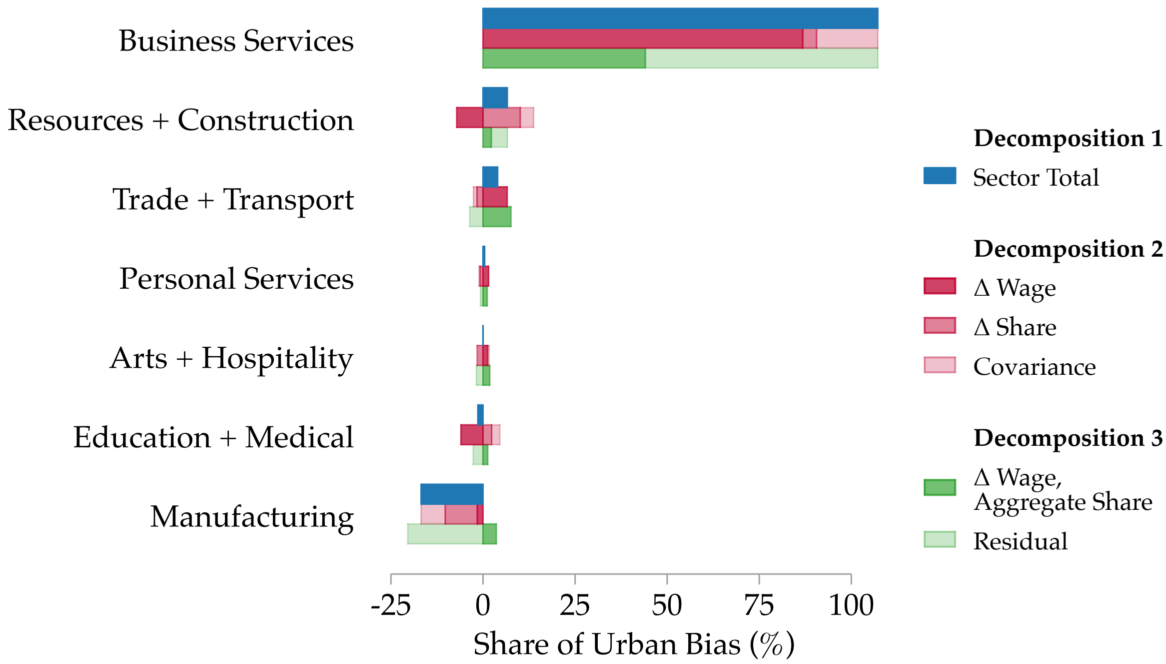
We apply the decomposition in equation (1) to study the contribution of each 1-digit NAICS sector to the wage-growth difference between the groups of commuting zones with above- and below-median density (cf. Figure 2). We use equation (1) to define the *share of urban-biased growth* accounted for by sector s, Ξ_s as follows:

$$(2) \qquad \Xi_s := \frac{\delta_{\ell' s} - \delta_{\ell s}}{g_{\ell'} - g_{\ell}},$$

where ℓ' refers to the above-median commuting zone group and ℓ to the below-median

[8]Figure OA.2 in the Online Appendix shows the figure with commuting zones ordered by their initial wage. Ordering cities based on their initial wage instead of population (density) yields different rankings. Many low-density cities in the US have high wages, and some high-density cities do not.

FIGURE 3: THE SECTORAL ORIGINS OF URBAN-BIASED GROWTH



Notes: The figure decomposes the difference in 1980-2015 wage growth between commuting zones with above-median and below-median densities in 1980 into the contributions of each NAICS-1 sector. The blue bars show the share of the wage growth difference accounted for by each sector (cf. equation (2)). The red bars decompose the blue bars into the separate contributions of within-industry wage growth, across industry relocation, and a covariance term (cf. equation (3)). The green bars decompose the blue bars into a component due to wage growth difference if all commuting zones had the same sectoral employment shares and a residual component (cf. equation (4)). The underlying data come from the US Census Bureau's Longitudinal Business Database and cover all US private, non-farm employer establishments. We classify above-median density commuting zones as the highest density commuting zones jointly accounting for 50% of 1980 employment. All values are adjusted for inflation to 2015 dollars using the BEA PCE price index.

commuting zone group.

The blue bars in Figure 3 show the share of urban-biased growth accounted for by each sector. The decomposition reveals that the Business Services sector accounted for most urban-biased wage growth.[9] Growth in all other sectors was remarkably balanced across high- and low-density commuting zones. The only sector that contributed negatively to urban-biased growth is Manufacturing (NAICS-3). Since manufacturing jobs were high-paying on average in 1980, their disproportionate disappearance from big cities lowered relative growth in the high-density cities.

The Business Services sector comprises industries often associated with high-population-density locations: Professional Services, Finance and Insurance, Management of Companies, Information, Administrative Services, and Real Estate Services. All 2-digit sub-industries of the Business Services sector experienced substantial urban-

[9]The NAICS-5 sector accounted for 20% of national employment in 1980, 66% of which were in high-population-density commuting zones. These numbers had changed to 26% and 61% by 2015.

biased growth.[10] In order of contribution, the industries contributing most to the sector's urban-biased growth are Professional Services, Finance, and Information. Wage growth in the "Management of Companies" sub-industry, which mainly captures large companies' headquarters establishments, has been strongly urban-biased. However, it accounts for little urban-biased growth simply because the sector is small relative to other subindustries of the Business Services sector.[11]

A sector's contribution to urban-biased growth can reflect wage growth or changes in its employment share. To differentiate these two channels, we decompose the contribution of each sector in equation (1) as follows:

$$\delta_{\ell s} = \overbrace{\underbrace{\frac{\mu_{s\ell t}\Delta w_{\ell st}}{\bar{w}_{\ell t}}}_{\Delta \text{ Wage}} + \underbrace{\frac{w_{\ell st}\Delta \mu_{s\ell t}}{\bar{w}_{\ell t}}}_{\Delta \text{ Share}} + \underbrace{\frac{\Delta \mu_{s\ell t}\Delta w_{\ell st}}{\bar{w}_{\ell t}}}_{\text{Covariance}}}^{\text{Decomposition 2}}. \tag{3}$$

Equation (3) divides the contribution of each sector into parts due to wage growth holding employment shares fixed, employment share changes holding wages fixed, and a covariance term. The red bars in Figure 3 use equation (3) to decompose a sector's contribution to urban-biased growth into these components. Urban-biased growth of the Business Services sector mainly reflects urban-biased wage growth *within* the sector; reallocation of employment across sectors plays only a minor role.

The wage-growth term in equation (3) combines cross-location differences in initial sectoral employment shares ("exposure") with cross-location differences in wage growth. To isolate cross-location differences in wage growth from differences in exposure, we extract a component from the wage term in equation (3) that holds exposure constant across locations:

$$\delta_{\ell st} = \overbrace{\underbrace{\frac{\mu_{st}\Delta w_{\ell st}}{\bar{w}_{\ell t}}}_{\Delta \text{ Wage, Aggregate Share}} + \underbrace{\xi_{\ell st}}_{\text{Residual}}}^{\text{Decomposition 3}}, \tag{4}$$

where μ_{st} is the employment share of sector s in the aggregate economy. We use equation (4) to decompose further the contribution to each sector's urban-biased growth in the green bars in Figure 3. We find that wage-growth differences within the Busi-

[10] See Figure OA.7 in the Online Appendix for the decomposition in equation (2) at the 2-digit NAICS level.

[11] Since our paper first appeared, headquarters establishments have recently received attention in the literature since advances in IT have changed how they control their associated production sites. Kleinman (2022) show headquarters and their associated establishments contribute to aggregate wage inequality in the US economy; Jiang (2023) studies how the expansion of their establishment networks affects the aggregate efficiency of the aggregate economy.

ness Services sector can account for almost 50% of urban-biased growth, even after controlling for exposure differences across locations.

Lastly, we note the Business Services sector has also experienced strong aggregate growth between 1980 and 2015. During that period, aggregate Business Services employment expanded faster than any other 1-digit NAICS sector, and aggregate wage growth was twice as fast as that of the second fastest-growing sector.[12]

In summary, urban-biased growth mainly occurred in Business Services, and reflected large within-sector differences in wage growth across locations.

Fact 2: Large establishments account for most urban-biased Business Services growth.

To understand the role of establishment size in contributing to urban-biased growth, we split all establishments into "large" (at least 100 employees) and "small" (less than 100 employees), where each group accounted for roughly 50% of US employment in 1980. We index these establishment "types" by e.

We modify the decomposition in equation (1) to decompose the contribution of each sector further into the contribution due to large and small establishments. The blue bars in Figure 4 show large Business Services establishments account for almost 70% of urban-biased growth. The blue bars are additive, so the large and small establishment components within the Business Services sector add to a sector's total contribution to urban-biased growth.[13]

The red bars in Figure 4 decompose each establishment type's contribution into wage growth versus employment-share growth. Most of the contribution of large Business Services establishments reflects wage growth rather than an increase in their local employment shares. The negative share component of other sectors' large establishments reflects the disproportionate decline of large, high-paying manufacturing establishments in big cities.

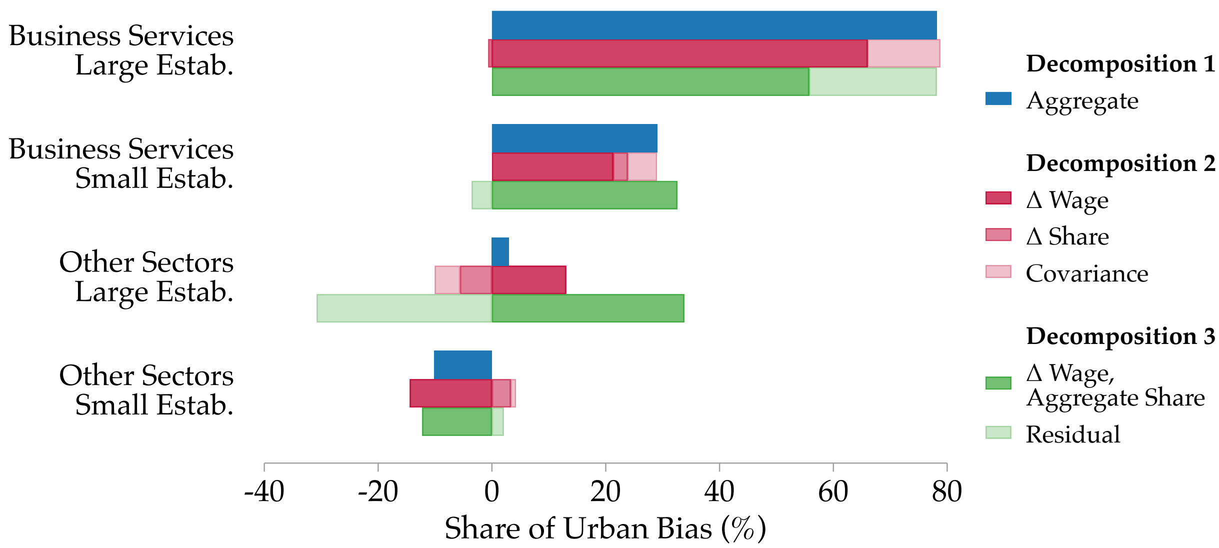
Next, we further decompose the contribution of the wage growth component. We again isolate a component that interacts local wage growth at large Business Services establishments with the employment share of such establishments in the aggregate economy (analogously to Figure 3 above). The dark green bar in Figure 4 shows the wage change alone, given identical exposure, accounts for the vast majority of the contribution of large Business Services establishments.

The finding that differential wage growth at large establishments is the central driver of

[12]Figure OA.12 in the Online Appendix shows aggregate wage and employment growth by sector.

[13]By construction, large and small establishments account for about 50% of total employment, so the outsized contribution of large establishments to urban-biased growth does not reflect sectoral size differences.

FIGURE 4: ESTABLISHMENT SIZE AND URBAN-BIASED GROWTH



Notes: The figure decomposes the difference in 1980-2015 wage growth between commuting zones with above-median and below-median densities in 1980 into the contributions of each NAICS-1 sector. The blue bars show the share of the wage growth difference accounted for by each sector and establishment type (cf. equation (2)). The red bars decompose the blue bars into the separate contributions of within-industry wage growth, across industry relocation, and a covariance term (cf. equation (3)). The green bars decompose the blue bars into a component due to wage growth difference if all commuting zones had the same sectoral employment shares and a residual component (cf. equation (4)). The underlying data come from the US Census Bureau's Longitudinal Business Database and cover all US private, non-farm employer establishments. We classify above-median density commuting zones as the highest density commuting zones jointly accounting for 50% of 1980 employment. We classify large establishments as the largest establishments jointly accounting for 50% of 1980 employment, leading to an employment cutoff for large firms of 108 employees. All values are adjusted for inflation to 2015 dollars using the BEA PCE price index.

urban-biased growth reflects two empirical regularities about employment at large Business Services establishments across commuting zones. First, while large establishments account for a larger share of local employment in big cities than small establishments, these employment share differences are minor relative to the vast differences in wage growth at large establishments across commuting zones. This finding helps understand why the dark green bars in Figure 4 that capture the wage growth contribution to urban-biased growth in a world with equalized large establishment employment shares across locations still explain most of the urban-biased growth in the data. Second, large Business Services establishments' employment shares across commuting zones have mostly stayed the same between 1980 and 2015; their employment shares in the highest density decile even fell slightly, explaining why the "share change" component in Figure 4 is mildly negative.[14]

In the Online Appendix, we show our findings in this section change little when we instead classify establishments into large and small based on the size of the firm

[14]Figure OA.9 in the Online Appendix shows wages and employment shares at large and small establishments across commuting zones these results.

that owns them. The establishments of large Business Services firms tend to be large themselves. As such, our theory focuses on establishments, and has little to say about their ownership structure.

In summary, large establishments are essential in accounting for the urban-biased growth of the US economy. Faster wage growth at large Business Services establishments in big cities explains most urban-biased growth.

Fact 3: IT investment is concentrated in large, urban Business Services firms.

Facts 1 and 2 showed wage growth at large establishments in the Professional Services, Finance, and Information industries accounted for most of the urban-biased growth in the data. These industries are often associated with the intensive use of IT, such as computers and software. In this section, we provide evidence that investments in IT capital occurred predominantly in high-density commuting zones and at large Business Services establishments, making the adoption of information technologies a candidate explanation for our findings in Facts 1 and 2 above.

We use the BEA Fixed-Asset Tables as a source of information on investments in different types of capital by sector. In the data, we define IT capital as all capital types falling into one of three subgroups: custom software, pre-packaged software, and hardware. We deflate the value of all investments by asset-specific deflators provided by the BEA.[15]

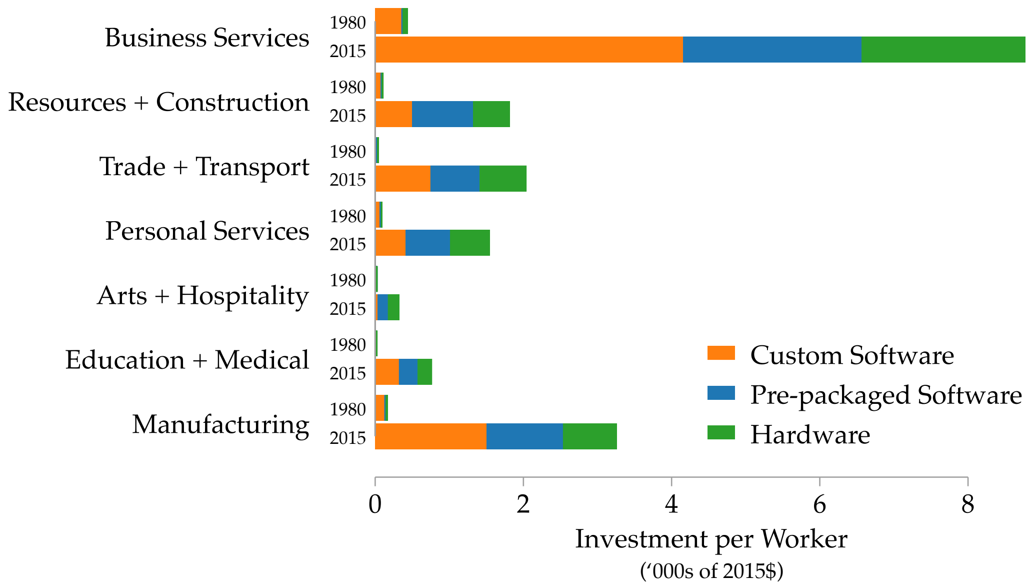
Figure 5 shows IT investments per worker in 1980 and 2015 for each NAICS-1 sector ordered by their contribution to urban-biased growth. In 1980, the Business Services sector already made higher IT investments per worker than any other sector. By 2015, the Business Services sector invested almost three times more than any other sector. The Business Services sector does not set itself apart regarding per-worker investments in any other type of capital besides IT; overall, it is not particularly capital-intensive.[16]

Next, we provide evidence that IT technology investments occurred predominantly in large firms in high-density commuting zones in the Business Services sector.

We use information from the 2013 Information and Communication Technology supplement to the US Census' Annual Capital Expenditure Survey (ACES) to disaggregate each sector's IT expenditures across commuting zones and production establishments. The ACES reports the capitalized and non-capitalized expenditures on various IT categories. A drawback of the ACES data is that they report expenditure at the firm rather than the establishment level. Firms with multiple establishments may have no unique sector or commuting zone. We merge the ACES data with the LBD data to observe the location, sector, payroll, and employment of each establishment associated with a

[15]See the Online Appendix for more details.
[16]The Online Appendix shows each sector's capital investments per worker in non-IT capital.

FIGURE 5: INFORMATION TECHNOLOGY INVESTMENTS PER WORKER ACROSS SECTORS



Notes: The figure shows investment per worker for different information technology assets across 1-digit NAICS sectors in 1980 and 2015. Data on capital investments in each sector are from the Bureau of Economic Activity. Data on employment in each sector are from the Quarterly Census of Employment and Wages. Proprietary software refers to BEA codes ENS2 and ENS3; pre-packaged software refers to ENS1, and hardware to EP1A-EP31. Sectors appear in order of their contribution to urban-biased growth. All values are adjusted using the BEA's asset-specific investment-price deflators to 2015 dollars.

firm. We measure the population density associated with multi-establishment firms as the average density across the commuting zones of all its establishments, weighted by each establishment's employment. We also define such firms' "Business Services employment share" as the fraction of their employment at establishments with a NAICS-5 code; the variable is one for single-establishment Business Services firms. For most observations, the Business Services employment share is either zero or one.

The first column of Table 1 show IT expenditure per worker also exhibited a strong urban bias. The second column adds controls for a firm's Business Services employment share. It shows the urban bias of IT investments was particularly strong among Business Services firms. For a firm with only Business Services employment, doubling log population density raises IT expenditure per capita by $806, as opposed to only $155 at a firm without any Business Services employment.

Next, we document the role of large firms and establishments in the urban bias of IT investments. Column 3 of Table 1 shows IT investments per worker increase in firm size. Our findings corroborate recent evidence by Lashkari et al. (2024), who documented similar facts in firm-level microdata from France. Column 4 shows the relationship between firm size and IT investment per capita was particularly strong in the Business Services sector.

Corroborating the evidence from Columns 1-4, Columns 5 and 6 show most IT in-

TABLE 1: IT EXPENDITURE, POPULATION DENSITY, AND ESTABLISHMENT SIZE

	IT Expenditure/Worker (x $1,000)					
	(1)	(2)	(3)	(4)	(5)	(6)
Log Population Density	0.469***	0.155***			0.00140	0.101*
	(0.0299)	(0.0224)			(0.0520)	(0.0442)
Log Employment			0.352***	0.181***	-0.170**	0.167***
			(0.0158)	(0.0132)	(0.0607)	(0.0450)
Log Emp. × 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)
× Log Density		0.651***				-0.182
		(0.0943)				(0.140)
× Log Emp.				0.539***		-0.456*
				(0.0452)		(0.198)
× Log Emp. × Log Density						0.163***
						(0.0346)
No. of Firms	45,000	45,000	45,000	45,000	45,000	45,000

Notes: Standard errors in parentheses. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. The table shows a regression of firm-level IT expenditure per employee (in thousands of 2013 dollars) on the log of a firm's average commuting zone population density, the log of the firm's employment size, and its Business Services employment share. For multi-establishment firms, the average commuting zone density is the employment-weighted average population density in 1980 across each establishment's commuting zone. The Business Services employment share is the share of a firm's employment at establishments with a NAICS-5 industry code. The data come from the 2013 Longitudinal Business Database and the ACES/ICTS survey. For disclosure reasons, the sample size is rounded to the closest thousand.

vestments per capita occurred at large Business Services firms in the highest-density commuting zones. After controlling for firm size and Business Services employment share, the density coefficient shrinks substantially relative to Column 1. This finding suggests firm-size differences across commuting zones and investment differences across the firm size distribution explain most of the aggregate relationship between density and IT investments.

The ACES is the only US Census data product that provides information on IT investments at service firms. However, it is only available in select recent years, and reports all information on the firm instead of the establishment level. In the Online Appendix, we corroborate the evidence in Table 1 with data purchased from Spiceworks, a commercial data provider, that are recorded at the establishment level.[17] Using the Spiceworks data, we replicate Table 1 and find quantitatively similar results. The data include detailed information on IT expenditure for a large sample of Business Services establishments. These data confirm IT expenditure per worker increases in location

[17] The Spiceworks data was formerly known as Ci Technology Database, produced by the Aberdeen Group, and before that as Harte-Hanks data. Due to the Spiceworks's broad coverage and high accuracy, many prior academic publications in economics have used it as a source of information (e.g., Bresnahan, Brynjolfsson, and Hitt, 2002; Beaudry et al., 2010; Bloom, Draca, and Van Reenen, 2016).

density and employment size for Business Services establishments, and much less so for other industries.

In summary, this section provided direct evidence that large Business Services firms in high-density commuting zones invested more heavily in IT capital than other sectors and firms. Our three facts above are consistent with the view that changes in the IT capital usage of Business Services establishments led to changes in their workforce composition and wage structure that gave rise to the urban-biased growth phenomenon.

1.4 The Role of Education

An extensive literature studies changes in the educational composition and skilled-wage structure of cities (Moretti, 2012; Diamond, 2016). Such changes could contribute to the urban-biased growth phenomenon in two ways. First, if more educated workers start moving to big cities, average wages in cities would increase because educated workers tend to earn higher average wages. Second, the contribution of the Business Services sector might reflect a general urban bias in the wage growth of skilled workers, combined with the fact that many of them work in the Business Services sector. This section studies both channels and shows they contribute little to urban-biased growth.

Since the LBD lacks demographic information on workers, we use data from the US Decennial Census and the American Community Survey (ACS) to study the role of education. Relative to the LBD, the Census contains information on individual workers' characteristics, such as education, but it is self-reported. We aggregate the wage and employment data to the commuting zone and sector level separately for workers with at least a college degree (”college”) and those with less education (”non-college”).[18]

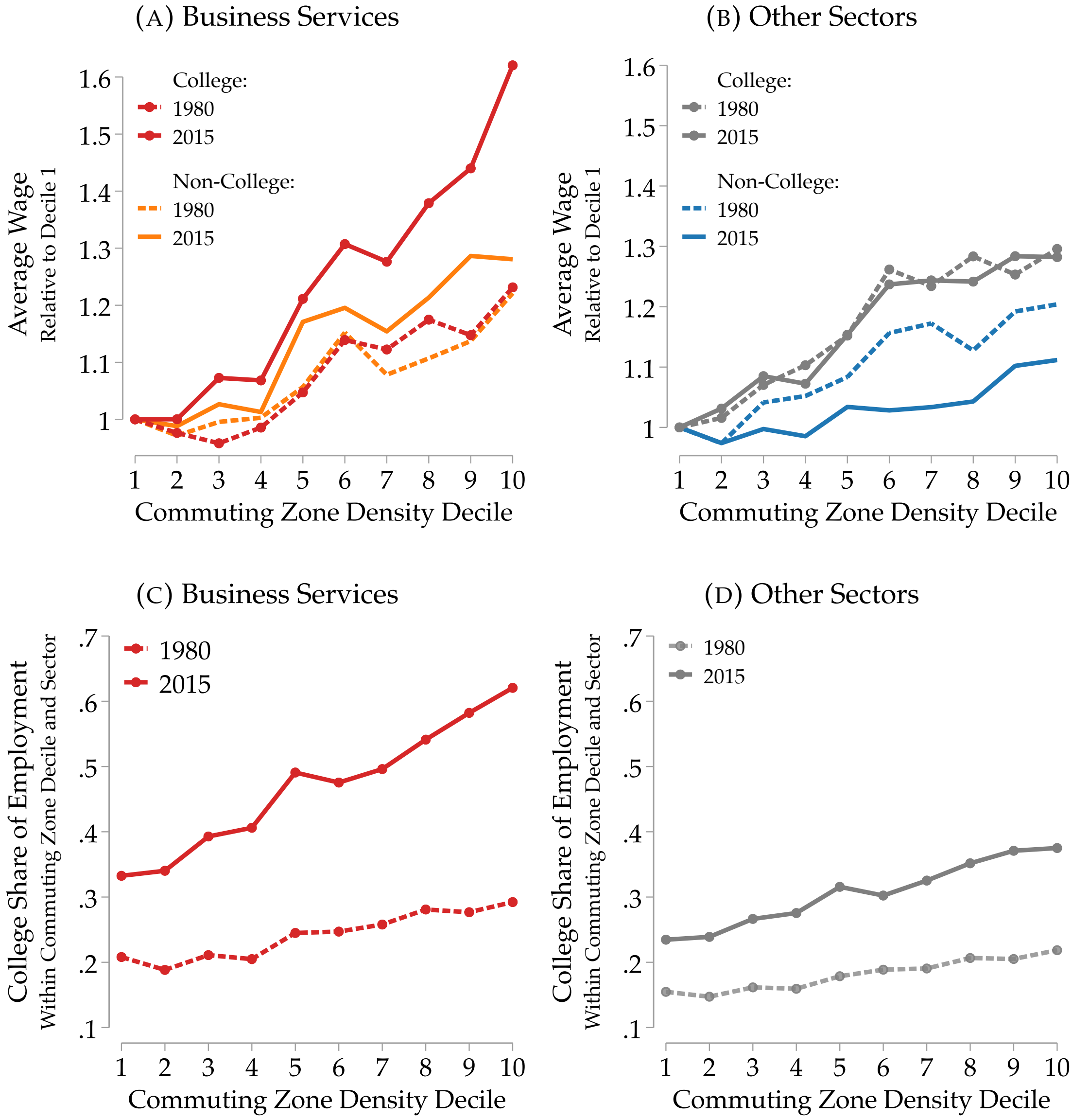
We begin by quantifying the role of changes in the composition of the urban workforce as a contributor to urban-biased growth. We decompose urban-biased growth into the contribution of the observed changes in the composition of each sector holding wages and sectoral employment shares fixed at their 1980 level and a residual term capturing changes in wages and employment shares.[19] We find that the shift of the Business Services workforce toward college-educated workers alone accounted for less than one-fifth of the sector's urban-biased growth. Across all sectors, the disproportionate shift toward college-educated workers in big cities can explain around 30% of all urban-biased growth, most of which reflects changes inside Business Services.

Compositional changes do not contribute much because the urban-biased increase in the college share of employment is moderate relative to spatial differences in wage growth within the Business Services sector. The bottom row of Figure 6 shows the college shares

[18]See the Supplementary Material for further details.

[19]Table OA.2 in the Online Appendix presents these results. In the Online Appendix, we also show the equations used for the decomposition.

FIGURE 6: THE ROLE OF EDUCATION IN URBAN-BIASED GROWTH



Notes: This figure shows average annual wages and college employment shares across commuting zones (Tolbert and Sizer, 1996) sorted into deciles of increasing population density, separately for Business Services and the rest of the economy in 1980 and 2015. Each decile accounts for one-tenth of the US population in 1980. The underlying data come from the 1980 US Decennial Census and the 2015 American Community Survey. Panel A shows average wages among college- and non-college-educated workers across commuting zone deciles in the Business Services sector between 1980 and 2015; Panel B shows the same outside the Business Services sector. Panel C shows college employment shares within the Business Services sector across commuting zone deciles in 1980 and 2015; Panel D shows the same outside the Business Services.

of employment within Business Services (left) and in the rest of the economy (right) across commuting zones and time. The college share of employment has increased more in big cities than in small cities, so Business Services now employs a larger fraction of high-skill workers. However, quantitatively, the urban-biased wage growth of both education types within the Business Services contributes much more to the urban-biased growth of the US economy.

Next, we study whether the urban-biased growth of the Business Services sector reflects a general urban bias in the wage growth of educated workers. The top row of Figure 6 shows wages for college- and non-college-educated workers in Business Services (left) and outside (right). In the Business Services sector, the wage-density gradient has increased over time for both types of workers, but more so for the skilled; conditional on working in Business Services, college- and non-college-educated workers have both experienced urban-biased growth. However, no urban-biased wage growth occurs outside the Business Services sector. The urban-wage gradient is stagnant for college-educated workers and decreasing for non-college-educated workers.[20] In other words, college-educated workers who did not work in Business Services experienced no urban-biased wage growth; such workers account for over 70% of all college-educated workers.

The preceding results validate the usefulness of our sectoral perspective by showing that urban-biased growth is a feature of the Business Services sector rather than a particular skill group.[21] At the same time, the hallmark features of skill-biased technical change, educational deepening and increases in the college wage premium, are particularly strong in the Business Services sector. These patterns align well with the urban-biased expenditure on IT we documented above, as the economics literature has long associated such technology with a distinct skill bias.

2. Theory

This section introduces a theory of uneven spatial growth through investment-specific technical change. The theory formalizes the relationship between capital use, factor prices, and firm size across locations in general equilibrium. It provides general conditions under which a spatially-neutral aggregate fall in investment prices can generate urban-biased growth.

2.1 Model Setup

The model consists of set of locations indexed by ℓ and sectors indexed by s. Production occurs using a set of factors $\mathcal{F}$, indexed by f. We differentiate three subsets of factors based on their mobility across locations and sectors. *Capital* factors are freely mobile across locations and sectors. *Labor* factors face relocation frictions across locations and sectors. *Commercially-zoned land* is immobile and specific to locations and sectors. We denote the corresponding subsets of $\mathcal{F}$ by $\mathcal{F}^K$, $\mathcal{F}^L$, and $\mathcal{F}^M$, respectively. Trade across

[20]David Autor discussed the decline of the urban-wage premium of non-college-educated workers in his Richard T. Ely Lecture in 2019 (see Autor, 2019).

[21]Other papers in the literature have studied the changes in the composition of big cities towards so-called cognitive non-routine ("CNR") occupations (Rossi-Hansberg, Sarte, and Schwartzman, 2019; Jaimovich and Siu, 2020). In the Online Appendix, we present a similar analysis as in this section for CNR occupations. Biased employment growth in CNR occupations (or 'occupational deepening') explains little of the urban bias in US growth.

locations is free. Time is discrete, indexed by t, and we suppress time subscripts where possible.

Production and Market Structure. The economy has a single final good that serves as the numeraire. To produce it, a representative firm combines varieties produced by individual firms i and sectors s using a nested CES aggregator with within-sector elasticity of substitution ι_s and across-sector elasticity γ.

An individual firm i is defined by its location and sector, its productivity, z_i, and the differentiated variety for which it owns a blueprint. Firm i's production technology is given by

$$y_i = z_i F_{\ell s}(y_i, \mathbf{x}), \tag{5}$$

where $\mathbf{x} = \{x_f\}$ is a vector of factor inputs. We assume the production function $F_{\ell s}(y_i, \cdot)$ is strictly positive, continuously differentiable, and increasing in the quantities of all inputs. Equation (5) allows for arbitrary productivity differences across locations and sectors. In addition, the production function is implicitly defined, allowing for the level of output y_i to affect the marginal products of factors.

We define two general elasticities that describe how a firm's optimal factor intensities change with factor prices and with its scale of production:

$$\sigma_{ff'} := \frac{\partial \log x_f / x_{f'}}{\partial \log w_{f'} / w_f} \quad \text{and} \quad \epsilon_{ff'} := \frac{\partial \log x_f / x_{f'}}{\partial \log y}, \tag{6}$$

where w_f denotes the unit rental rate of factor f. The term $\sigma_{ff'}$ denotes the elasticity of substitution between two factors f and f'; it takes a constant value in the CES case and is 1 in the Cobb-Douglas case. We refer to term $\epsilon_{ff'}$ as the "scale elasticity." The scale elasticity captures how the optimal input ratio of two factors changes with a firm's level of output, or *scale*. For homothetic production functions $\epsilon_{ff'} = 0$ for any two factors f and f'. We allow for $\epsilon_{ff'} \neq 0$ to capture the empirical relationship between firm size and capital intensity presented in Fact 3. Both elasticities are functions of technologies and factors prices and can hence differ across locations and sectors.

For every level of output, firms choose factor inputs to minimize their variable production costs given local factor prices and technologies. We denote the variable cost function of firm i in a location-sector by

$$c_{\ell s}(z_i; y_i, \mathbf{w}_{\ell s}) = y_i z_i^{-1} v_{\ell s}(y_i, \mathbf{w}_{\ell s}),$$

where $\mathbf{w}_{\ell s} = \{w_{\ell s f}\}$ denotes the vector of factor prices for all factors. The unit variable cost, $z_i^{-1} v_{\ell s}(y_i, \mathbf{w}_{\ell s})$, varies with firm scale if $\epsilon_{ff'} \neq 0$, that is, as long as the technology is a non-homothetic function.

The representative firm's profit maximization implies firm i's revenue function is $y_i^{\zeta_s}\mathcal{D}_s$, where $\mathcal{D}_s$ is an endogenous measure of aggregate sectoral demand, and the demand elasticity is a function of the elasticity of substitution across firm varieties within a sector, $\zeta_s = (\iota_s - 1)/\iota_s$.

Each period, firms choose a level of output to maximize variable profits given their cost and revenue functions:

$$\pi_{\ell s}(z_i) := \max_y \left[y^{\zeta_s}\mathcal{D}_s - c_{\ell s}(z_i; y, \mathbf{w}_{\ell s}) \right].$$

To enter a location-sector pair, firms incur a fixed entry cost. The entry cost is produced using a technology $g_{\ell s}(\mathbf{x})$, where $g_{\ell s}(\cdot)$ is strictly positive, continuously differentiable, increasing in each argument, and homogeneous of degree 1. We denote the corresponding entry-cost function by $e_{\ell s}(\mathbf{w}_{\ell s})$.

Upon entry, firms draw their productivity $z_i \in (0, \infty)$ from a distribution $\Omega_s(z)$. We assume entry costs are sunk to rule out selection on entry, in line with the empirical evidence in Combes, Duranton, Gobillon, Puga, and Roux (2012). Firms exit at an exogenous rate ξ, consistent with evidence from Walsh (2023).

In each period, new firms enter a given location-sector as long as the present discounted value of expected profits exceeds the entry costs, resulting in the following free-entry condition:

$$e_{\ell s}(\mathbf{w}_{\ell st}) = \int V_{\ell st}(z) d\Omega_s(z),$$

where $V_{\ell st}(z)$ is the present discounted value of a firm with efficiency z in location ℓ and sector s at time t. We denote the total number of active firms in a given location-sector at time t by $N_{\ell st}$, which combines new entrants and surviving incumbents.

Local Factor Supply. We partition the vector of location-sector factor prices and factor supplies into subvectors for each factor type, capital, labor, and commercial land:

$$\mathbf{w}_{\ell s} = \{\mathbf{w}_{\ell s}^K, \mathbf{w}_{\ell s}^L, \mathbf{w}_{\ell s}^M\} \quad \text{and} \quad \mathbf{X}_{\ell s} = \{\mathbf{X}_{\ell s}^K, \mathbf{X}_{\ell s}^L, \mathbf{X}_{\ell s}^M\},$$

where $\mathbf{w}_{\ell s}^K = \{w_{\ell sf}^K\}$ and $\mathbf{X}_{\ell s}^K = \{X_{\ell sf}^K\}$, and similarly for labor and commercial land. We refer to $\mathbf{w}_{\ell s}^K$ as the rental rates of capital, to $\mathbf{w}_{\ell s}^L$ as the wages of different types of workers, and to $\mathbf{w}_{\ell s}^M$ as rents for commercial land.

The three types of factors differ in their mobility across locations and sectors. *Capital* factors $f \in \mathcal{F}^K$ are freely mobile within the economy, so that their rental rates are the same across locations and sectors. The total national supply of capital, $\mathbf{X}^K$, is endogenous and described below.

Commercially-zoned land-type factors $f \in \mathcal{F}^M$ are immobile and non-tradable, so that rental rates $\mathbf{w}^M_{\ell s}$ differ across location-sector pairs. The supply of commercially-zoned land within each location-sector, $\mathbf{X}^M_{\ell s}$, is exogenous.

Labor factors are imperfectly mobile across locations; their national supply, $\mathbf{X}^L$, is exogenous, but their local supply, $\mathbf{X}^L_{\ell s}$, depends on the utility-maximizing choices of individuals. Each period, an individual worker j of labor type f chooses a location, sector, and quantities of residential land (o) and the final good (c) to solve the following utility-maximization problem:

$$\max_{\ell}\{v^j_\ell \mathbb{E}_{v_s} \max_{s,o,c}\{o^{\alpha_f} c^{1-\alpha_f} v^j_s\}\} \quad \text{subject to} \quad r_\ell o + c \leq w^L_{\ell s f}$$

where v^j_ℓ and v^j_s are idiosyncratic preference shocks for sectors and locations, α_f is the expenditure share of type f workers on residential land, and r_ℓ is the rental rate of residential land. Workers first learn their location-specific shocks and only learn about their sectoral preferences upon arriving in a location. The expectation operator indicates workers have to form expectations over their sector-specific shocks within a location when making their location decisions.

We assume workers draw their idiosyncratic preference shocks for each location and sector from separate Fréchet distributions with inverse scale parameters $B_{\ell f}$ and $B_{\ell s f}$ and shape parameters ϱ^1_f and ϱ^2_f. These assumptions yield familiar expressions for the fraction of type-f workers choosing to live in location ℓ, $\lambda_{\ell f}$, and for the fraction of type-f workers in location ℓ choosing to work in sector s, $\mu_{\ell s f}$:

$$\lambda_{\ell f} = \frac{B_{\ell f}(r_\ell^{-\alpha_f}\Psi_{\ell f})^{\varrho^1_f}}{\sum_\ell B_{\ell f}(r_\ell^{-\alpha_f}\Psi_{\ell f})^{\varrho^1_f}} \quad \text{and} \quad \mu_{\ell s f} = \frac{B_{\ell s f}(w^L_{\ell s f})^{\varrho^2_f}}{\sum_s B_{\ell s f}(w^L_{\ell s f})^{\varrho^2_f}}, \tag{7}$$

where $\Psi_{\ell f} := (\sum_s B_{\ell s f}(w^L_{\ell s f})^{\varrho^2_f})^{1/\varrho^2_f}$ is the expected utility of a type-f worker in location ℓ prior to learning their sectoral preference shocks. The terms $B_{\ell f}$ and $B_{\ell s f}$ play the role of type-specific amenity terms for locations and sectors. The local labor supply of type-f labor in a given location-sector is $X^L_{\ell s f} = \lambda_{\ell f}\mu_{\ell s f} X^L_f$.

Investment Decisions and Factor Ownership. A unit mass of identical atomistic capitalists makes all dynamic investment decisions in the economy. They own all firms, capital, commercially-zoned land, and residential land.

In each period, capitalists decide how much of the final consumption good to consume, how much to invest in each type of capital, and how many firms to create in each location and sector, to maximize the following utility function:

$$\max_{\{C_t\},\{\mathbf{X}^K_{t+1}\},\{N_{\ell st+1}\}} \sum_{t=0}^{\infty} \beta^t \log(C_t) \tag{8}$$

subject to a set of period budget constraints:

$$\begin{aligned} C_t + \mathbf{p}^K_t(\mathbf{X}^K_{t+1} - (1-\delta^K_t)\mathbf{X}^K_t) + \sum_{\ell,s} e_{\ell s}(\mathbf{w}_{\ell st})(N_{\ell st+1} - (1-\xi)N_{\ell st}) \\ = \mathbf{w}^K_t \mathbf{X}^K_t + \sum_{\ell,s} \mathbf{w}^M_{\ell st}\mathbf{X}^M_{\ell s} + \sum_{\ell,s} \Pi_{\ell st} + \sum_{\ell} r_{\ell t} O_\ell, \end{aligned}$$

where vector products are understood as dot products. The term $\beta \in (0,1)$ is the capitalists' discount rate, O_ℓ is the stock of residential land in location ℓ, and $\Pi_{\ell st}$ are total variable firm profits in a location-sector pair. The term $\mathbf{p}^K_t = \{p^K_{ft}\}$ is the vector of investment prices for the different types of capital and δ^K_t is the vector of capital depreciation rates, which may vary over time. Commercially-zoned land and residential land are in fixed supply.

A representative capital-producing firm transforms the final good into capital at capital-type-specific rates $\mathcal{Z}_{ft}$ so that $p^K_{ft} = 1/\mathcal{Z}_{ft}$. Period 0 has an initial supply of type f capital K_{f0}, and an initial stock of firms in each location and sector $N_{\ell s0}$. Finally, there is a set of non-negativity constraints on each asset.

Equilibrium. An equilibrium is a set of factor prices $\{\mathbf{w}_{\ell st}\}$, rental rates of residential land $\{r_{\ell t}\}$, investment prices of capital $\mathbf{p}^K_t$, consumption of capitalists C_t, a vector of aggregate capital stocks $\{\mathbf{X}^K_t\}$, number of firms in each location $\{N_{\ell st}\}$, and local labor supply $\{\mathbf{X}^L_{\ell st}\}$, such that in each period, (i) the capitalist solves the problem in equation (8), (ii) worker location decisions satisfy the expressions in equation (7), (iii) the labor and commercially-zoned land markets clear in every location and sector, for every type, (iv) the investment and spot capital markets clear for every capital type, (v) the market for residential land clears in every location, and (vi) the final-good market clears nationally.

A *steady-state equilibrium* is one in which all prices and allocations are constant across periods t.

2.2 Investment-Specific Technical Change and Urban-Biased Growth

In this section, we describe the effect of a decline in the investment price of capital on average wages across location-sector pairs in the steady state of the model. Any change in the investment price of type-f capital is the result of investment-specific technical change in $\mathcal{Z}_{ft}$, the rate at which the economy can transform the final good into type f capital. We take these productivity changes to be exogenous throughout the paper, representing fundamental technical progress in production of certain types of capital.

We define a firm's expected lifetime payments to factor f as

$$\Phi_{\ell sf} := \frac{\partial e_{\ell s}(\mathbf{w}_{\ell s})}{\partial w_{\ell sf}} w_{\ell sf} + \kappa \int \frac{\partial c_{\ell s}(z; y, \mathbf{w}_{\ell s})}{\partial w_{\ell sf}} w_{\ell sf} d\Omega_s(z),$$

which is the expected present-discounted value of a firm's payments to factor f over its lifetime in the steady state.[22] The term $\kappa := (\beta + 1)/(\beta\xi + 1) > 1$ is a combination of the capitalists' discount rate and the firm exit probability. By $\Phi^K_{\ell s} := \sum_{f \in \mathcal{F}^K} \Phi_{\ell sf}$, we denote the total payments to all capital types; we define similar terms for labor and commercial land. We also define the change in the average price of local factors (labor and commercial land) in a location-sector pair as

$$d \log \bar{w}_{\ell s} := \sum_{f \in \mathcal{F}^L \cup \mathcal{F}^M} \phi_{\ell sf} d \log w_{\ell sf} \quad \text{where} \quad \phi_{\ell sf} := \frac{\Phi_{\ell sf}}{\sum_{f' \in \mathcal{F}^L \cup \mathcal{F}^M} \Phi_{\ell sf'}},$$

which is the cost-share-weighted average rental rate growth across all local factors.

With this notation in hand, we establish the following theorem:

Theorem 1. *In the steady state, the general equilibrium response of average local-factor prices in a location-sector to a change in the investment price of type-f capital, p^K_f, is given by*

$$d \log \bar{w}_{\ell s} = -\frac{\Phi^K_{\ell sf}}{\Phi^L_{\ell s} + \Phi^M_{\ell s}} d \log p^K_f + \frac{\Phi^L_{\ell s} + \Phi^M_{\ell s} + \Phi^K_{\ell s}}{\Phi^L_{\ell s} + \Phi^M_{\ell s}} d \log \mathcal{D}_s.$$

Theorem 1 is the result of totally differentiating the free-entry condition. It shows the general equilibrium response of *average* local-factor prices to an exogenous decline in the investment price of capital is governed solely by relative factor payments in the location-sector. The details of production functions, firm heterogeneity, and factor supply are irrelevant. In particular, places with a greater ratio of payments to type f capital relative to local factors (land and labor) require a greater equilibrium response of local-factor prices.

The critical insight behind Theorem 1 is that the average price of local factors is pinned down by the free-entry condition alone, independently of factor supply curves. To see why, recall that the free-entry condition equates a firm's present discounted value of variable profits with the entry cost. Since capital rental rates do not vary across locations, the average rental rate of local factors has to adjust to offset variation in firm profitability induced by technological differences across locations and sectors. In particular, rental rates of local factors have to be higher in more productive locations

[22]The definition invokes Shephard's lemma on the entry-cost and unit-cost functions of the firm.

and sectors in equilibrium. Of course, the rental rate of any particular local factor also depends on its supply elasticity.

Now consider the effect of a decrease in the investment price of type f capital. Cheaper capital raises the variable profit of firms everywhere by lowering the unit rental rate of type-f capital. However, the extent to which variable profits rise differs across locations and depends on the importance of type-f capital in the cost structure of the average local firm. As a result, the *average* price of local factors has to rise differentially across location-sector pairs to restore free entry.[23]

Theorem 1 shows the effect of a change in the investment price of capital has two parts. The first term on the right-hand side is a *direct effect*. The higher the payments to type-f capital relative to payments to local factors in a location-sector, the more a falling capital price increases firm profitability. For given payments to local factors, the higher the payments to capital, the more significant the cost savings from a decline in its price, and the higher the average price of local factors has to rise to make up for these profitability gains. For given payments to capital, the lower the payments to local factors, the higher their price has to rise to achieve the same reduction in profitability.

The second term on the right-hand side of the equation in Theorem 1 presents an *indirect effect*. A decrease in the price of type-f capital also raises aggregate demand, which increases the sales of all firms. The intuition for the exposure of a location-sector to these changes is similar to the direct effect. All else equal, the higher a location's total factor payments, the larger its share of the aggregate economy and, hence, the more it is affected by changes in aggregate demand. Given total factor payments, the lower the payments to local factors, the more the average price of local factors has to adjust to restore free entry.

To build intuition for the cross-sectional implications of Theorem 1, we consider the following special case:

Corollary 1. *Consider a version of the economy with $\kappa \to 1$ and two factors of production, capital and labor. In this case, Theorem 1 reduces to:*

$$d\log w_{\ell s}^{L} = -\Lambda_{\ell s} d\log p^{K} + (1+\Lambda_{\ell s}) d\log \mathcal{D}_{s} \quad \textit{where} \quad \Lambda_{\ell s} := \frac{w^{K} X_{\ell s}^{K}}{w_{\ell s}^{L} X_{\ell s}^{L}}.$$

Corollary 1 shows in the two-factor version with a static entry decision, the cross-sectional variation in the wage response is summarized by a single exposure term $\Lambda_{\ell s}$, the ratio of total payments to capital relative to labor among all firms in location ℓ and sector s.

[23] Prices of local factors rise via firms increasing their output and more firms entering, both of which raise factor demand.

For a decline in capital investment prices to generate urban-biased growth, the exposure term has to increase with population density in the cross-section of locations. In the following subsection, we present a simple version of the model without firm heterogeneity to illustrate the determinants of the cross-sectional variation in exposure.

2.3 The Determinants of Exposure: A Simple Example

Consider a single-sector version of the model with just one type of capital, one type of labor, and no commercial land, in which $\kappa = 1$. We specialize the production function for variable and entry costs as follows:

$$y_i = A_\ell F(y_i, \mathbf{x}) \quad \text{and} \quad g(\mathbf{x}) = 1,$$

so that firms are homogeneous with $z = 1$ and locations only differ in a factor-neutral productivity shifter A_ℓ. We define urban-biased growth as faster wage growth in locations with higher location productivity, because higher population density is empirically strongly associated with higher labor productivity (see Ahlfeldt and Pietrostefani, 2019).[24] For simplicity, we restrict the non-homotheticity to generate a non-zero scale elasticity without inducing increasing returns, that is we assume $\epsilon_{KL} \neq 0$ and $\partial v(y, \mathbf{w}_\ell)/\partial y = 0$.

In this version of the model, Theorem 1 takes the form given in Corollary 1, so that the term Λ_ℓ is a sufficient statistic for exposure differences to investment-specific technical change across locations. Declines in the investment price of capital lead to urban-biased growth if the exposure elasticity is higher for locations with higher productivity. To see under what circumstances exposure is higher in more productive locations, we totally differentiate the expression for the exposure term in Corollary 1 in the cross-section of locations. The following expression shows how exposure changes as we move from less to more productive locations:

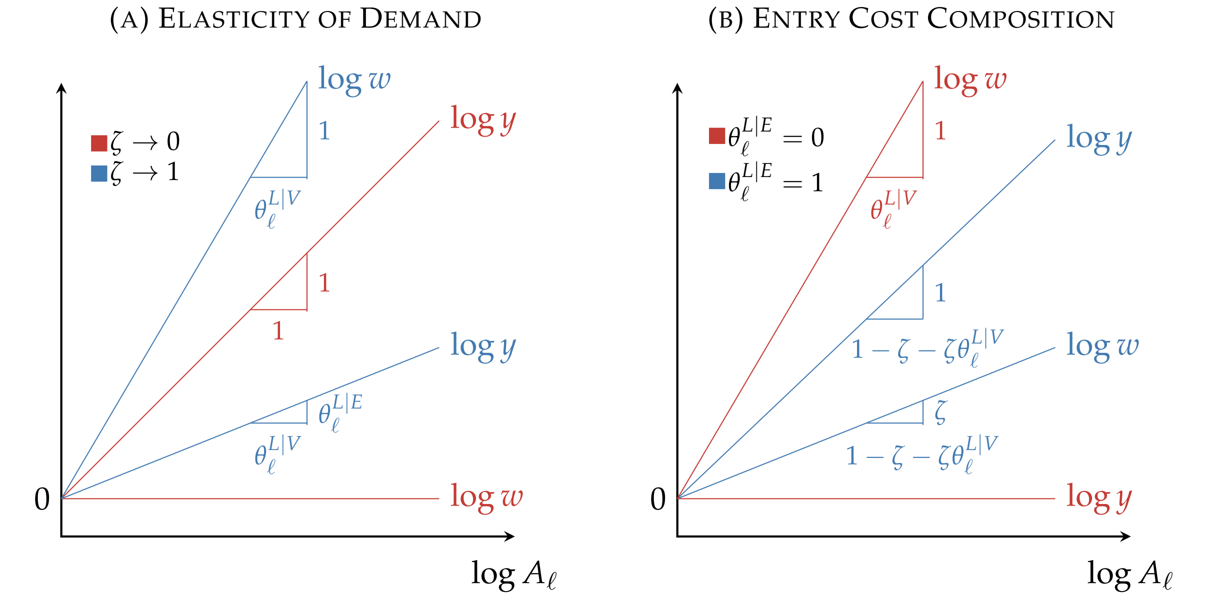
$$(9) \qquad \frac{d \log \Lambda_\ell}{d \log A_\ell} = \overbrace{(\sigma_{KL} - 1) \frac{d \log w_\ell^L}{d \log A_\ell}}^{\text{Neoclassical}} + \overbrace{\underbrace{\epsilon_{KL} \frac{d \log y}{d \log A_\ell}}_{\text{Non-Homotheticity}} + \underbrace{(\theta_\ell^{V|K} - \theta_\ell^{V|L}) \frac{d \log y / A_\ell}{d \log A_\ell}}_{\text{Entry vs. Variable Cost}}}^{\text{Scale}},$$

where $\theta_\ell^{V|K} \in [0, 1]$ and $\theta_\ell^{V|L} \in [0, 1]$ are the variable cost shares in total payments to capital and labor in location ℓ.

Equation (9) shows that two different channels determine the cross-sectional variation of exposure. The *neoclassical channel* captures the classic price and substitution effects

[24]Whether this will be true in the equilibrium of the model depends on the correlation of amenities for labor $B_{\ell s}$ with productivity A_ℓ. When we take the model to the data, we find that larger places indeed have higher productivities.

FIGURE 7: FIRM SCALE, WAGES, AND LOCAL PRODUCTIVITY



Notes: The figure shows how wages and firm output vary with productivity in the cross-section of locations in the simple version of the model. Higher values on the x-axis imply higher productivity and higher values on the y-axis imply higher wages or output. The left panel shows these relationships for two special cases of the demand elasticity for firm-specific varieties: Cobb-Douglas ($\zeta \to 0$) and perfect substitutes ($\zeta \to 1$). The right panel shows the same relationships for two special cases of the entry cost: no labor in the entry cost ($\theta_\ell^{L|E} = 0$) and only labor in the entry cost ($\theta_\ell^{L|E} = 1$).

in response to the higher wages associated with more productive locations. The price effect causes the ratio of capital to labor payments (i.e., the exposure elasticity) to fall as one moves from lower to higher-wage locations. The substitution effect reflects a shift of the cost structure towards capital whose price is not increasing and raises the ratio. The neoclassical channel implies that locations with higher productivity are less exposed to capital price changes as long as capital and labor are complements in production, that is $\sigma_{KL} < 1$.

We refer to the second part of equation (9) as the *scale channel*. The scale channel captures cross-sectional variation in the exposure elasticity resulting from the correlation of productivity and firm scale across locations. The scale channel has two components, both related to how firms' cost structures change with their scale. The first depends on how the optimal capital-labor ratio in variable production changes with firm size as captured by the scale elasticity ϵ_{KL}. Suppose that larger firms are more capital-intensive, so that $\epsilon_{KL} > 0$, and firms in high-productivity locations are larger on average. In that case, this channel pushes for a higher exposure statistic in productive locations.

The second component depends on how increasing output changes the loading on variable cost versus entry cost. If output increases faster than productivity, firms' total costs in high-productivity locations consist of a larger share of variable costs than those in low-productivity locations. This compositional difference raises exposure in more productive locations if variable costs are more capital intensive than entry costs

($\theta_\ell^{V|K} > \theta_\ell^{V|L}$).

Equation (9) also shows that how wages and firm scale vary with local productivity is important for the cross-sectional exposure patterns. The simple model permits explicit expressions describing wage and firm scale patterns in the cross-section of locations:

$$(10) \qquad \frac{d\log w_\ell^L}{d\log A_\ell} = \frac{\zeta}{\zeta\theta_\ell^{L|V} + (1-\zeta)\theta_\ell^{L|E}} \quad \text{and} \quad \frac{d\log y}{d\log A_\ell} = \frac{\theta_\ell^{L|E}}{\zeta\theta_\ell^{L|V} + (1-\zeta)\theta_\ell^{L|E}},$$

where $\theta_\ell^{L|V} \in [0,1]$ is labor share in variable costs, $\theta_\ell^{L|E} \in [0,1]$ is the labor share in entry costs, and $\zeta \in (0,1)$. The elasticity of demand and the share of the local factor (labor) in the entry cost play a crucial role in shaping the cross-sectional variation of wages, output, and location productivity. Figure 7 shows the expressions in equation (10) in two special cases that illustrate the role of the demand elasticity and the composition of the entry cost. The left panel of Figure 7 shows the Cobb-Douglas limit ($\zeta \to 0$) in red and the perfect-substitutes limit ($\zeta \to 1$) in blue. The right panel of Figure 7 shows the case without labor in the entry cost in red and with only labor in the entry cost in blue.

Figure 7 offers two takeaways essential for the rest of our analysis. First, except for the Cobb-Douglas limit case, wages always increase in local productivity.[25] As a result, the neoclassical channel is always active, pushing for lower exposure to capital price movements in more productive locations. Second, for firm scale to increase with productivity (and hence for the scale channel to be active), entry costs must rely on the local factor. The more important the local factor in entry cost, the more scale increases with local productivity, and the stronger the scale channel becomes.[26] The Online Appendix provides detailed intuition for the patterns in Figure 7.

Much of the intuition from the simple model carries over to our general theory. Consider the empirically relevant case for the Business Services sector in which more productive locations have higher wages, higher population density, and larger firms. In this case, the scale channel raises the exposure of high-density locations since larger firms tend to use IT capital more intensively. At the same time, the neoclassical channel implies high wages lower exposure as long as capital and labor are complements. Whether the scale channel dominates the neoclassical channel is a quantitative question.

The simple theory also highlights why urban-biased growth may be limited to some sectors. For sectors where firm size does not increase with population density, or capital

[25]In the Cobb-Douglas limit, each firm's revenue is a fixed fraction of national sales. Higher location productivity is one-for-one offset by lower prices for the firm's product so that the marginal product of labor is constant across locations.

[26]More productive locations have higher wages, so entry costs are higher in more productive locations if labor is in the entry cost. As a result, firms need to operate at a larger scale to make enough variable profits to pay for entry.

intensity does not increase with firm size, the neoclassical channel suggests declines in the rental rate of capital lead to rural-biased growth. Finally, more generally, for sectors that are not intensive users of capital, declines in investment prices do not lead to significant general equilibrium wage responses in any location.

2.4 Parameterization

To bring the model to the data, we need to specify the production function, the entry-cost function, and the distribution of firm heterogeneity.

We select a specific functional form for the production function in equation (5). We choose the canonical functional form to describe production functions with capital and low- and high-skill labor introduced in Krusell et al. (2000), but modify it to allow for a non-homotheticity. In particular, we specialize the production function in equation (5) as follows:

$$(11) \qquad y := z_i \left[\left(y^{\frac{\bar{\epsilon}_s}{\sigma_s}} (A^h_{\ell s})^{\frac{1}{\sigma_s}} h^{\frac{\sigma_s - 1}{\sigma_s}} + (A^k_s)^{\frac{1}{\sigma_s}} k^{\frac{\sigma_s - 1}{\sigma_s}} \right)^{\frac{\sigma_s}{\sigma_s - 1} \frac{\varphi_s - 1}{\varphi_s}} + (A^l_{\ell s})^{\frac{1}{\sigma_s}} l^{\frac{\varphi_s - 1}{\varphi_s}} \right]^{\frac{\varphi_s}{\varphi_s - 1}},$$

where σ_s denotes the elasticity of substitution between high-skill labor and capital, and φ_s denotes the elasticity of substitution between capital (or high-skill labor) and low-skill labor. $A^h_{\ell s}$ and $A^l_{\ell s}$ are location-sector-specific productivity shifters for high- and low-skill labor, and A^k_s is a sector-specific productivity shifter for capital.[27] Our quantitative analysis interprets k in equation (11) as IT capital. IT capital has experienced much more dramatic investment-specific technical change than other types of capital and appears as the most essential capital input in the Business Services sector.[28] Incorporating other equipment capital is straightforward, but our analysis absorbs them in the residual productivity terms for parsimony.

Equation (11) is part of a class of non-homothetic CES production functions. Sato (1977) introduced non-homothetic CES functions, and they have recently been used in a demand context by Comin et al. (2021).[29] If $\bar{\epsilon}_s = 0$, the production technology collapses to that in Krusell et al. (2000), in which each factor's marginal product is independent of the scale of production, and all firms in a location-sector have the same factor shares. If $\bar{\epsilon}_s \neq 0$, the value of a firm's marginal product of high-skill labor depends on its level of output, y.[30]

[27]Making these productivity shifters endogenous functions of local population size and composition, as in the urban literature, changes nothing fundamental about our exercise, and we do this in a robustness exercise below.

[28]See Figure OA.13 in the Online Appendix; other kinds of capital have seen mild price falls at best.

[29]To the best of our knowledge, Lashkari et al. (2024) and Trottner (2019) were the first to consider non-homothetic production functions in structural macro models.

[30]With the non-homothetic CES function, the elasticity of substitution continues to be constant at

The non-homotheticity parameter $\bar{\epsilon}_s$ is important in our theory, since it is a key ingredient in the scale channel outlined above. Given the functional form for the production function in equation (11), the general elasticities in equation (6) take more specific forms:

$$(12) \quad \epsilon_{kh} = -\bar{\epsilon}_s \qquad \epsilon_{kl} = -\bar{\epsilon}_s \frac{\varphi_s - \sigma_s}{1-\sigma_s}\theta_{\ell s}(w_{\ell s}, y) \qquad \epsilon_{hl} = \bar{\epsilon}_s[1 - \frac{\varphi_s - \sigma_s}{1-\sigma_s}]\theta_{\ell s}(w_{\ell s}, y),$$

where $\theta_{\ell s}(w_{\ell s}, y) \in (0, 1)$ is the cost share of skilled labor in the capital-skill bundle. The empirically relevant case is when $\varphi_s > 1 > \sigma_s > 0 > \bar{\epsilon}_s$. In this case, high-skill labor and capital are complements, whereas low-skill labor and capital are substitutes as in Krusell et al. (2000). In addition capital per (high- and low-skill) worker increases with firm size, in line with the evidence in Table 1. The expressions in equation (12) also justify the particular location of the non-homotheticity parameter in equation (11). Putting the non-homotheticity on high-skill labor is the only choice that can produce the positive relationship between capital- and skill-intensity and firm size seen in the data.

Importantly, the marginal rate of technical substitution between high-skill labor and capital implied by the production technology in equation (11) is given by:

$$\frac{dy/dh}{dy/dk} = (\frac{k}{h}\frac{A^h_{\ell s}}{A^k_s})^{\frac{1}{\sigma_s}} y^{\frac{\bar{\epsilon}_s}{\sigma_s}}$$

As long as $\bar{\epsilon}_s < 0$, the marginal rate of substitution is decreasing in firm output. In other words, capital and high-skill labor are more complementary at firms operating at a larger scale.[31] In line with this intuition, the non-homothetic CES production function can be micro-founded with firms choosing among a continuous menu of homothetic technologies that differ in their fixed costs and their factor intensities. Larger firms choose different technologies than smaller firms, as their larger scale means higher fixed costs and lower marginal cost technologies are profitable for them. See Trottner (2019) and Lashkari et al. (2024) for this and several other microfoundations.[32]

In addition, it is worth noting that $\bar{\epsilon}_s < 0$ implies increasing returns to scale in production. If $\sigma_s > 1$, a parameter restriction on the curvature of demand is necessary to ensure a firm's output choice is well defined. However, if $\sigma_s < 1$, marginal costs approach a constant in the limit as output grows, and no restriction on the curvature of demand is necessary. In the calibration section below, $\sigma_s < 1$ is the empirically relevant case.

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).

[31]This implies, for example, 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.

[32]Alternatively, other papers in the literature simply assume firm productivity is biased so that larger, more productive firms produce in a more capital- or skill-intensive way, see for example Burstein and Vogel (2017). The non-homothetic CES is a way of characterizing this idea.

Another critical aspect of equation (11) is that given structural parameters, we can choose its productivity shifters to match the observed data on average wages and total employment for each location-sector in each period. Similarly, the shifter on capital allows us to match the relative payments of capital to labor in each sector and at each point in time. By choosing the productivity shifters in this way, we can flexibly account for other sources of wage growth across location-sectors besides the one highlighted by our theory. Hence, our theoretical framework is helpful as an accounting device to understand which part of the wage growth in the data is due to changes in the investment price of capital relative to other sources of spatial growth.

We choose the following functional form for the entry-cost production function:

$$1 = g_s(\mathbf{x}) := \tau_s h^{\eta_s} l^{\eta_s} m_s^{1-2\eta_s},$$

where h, l denote high- and low-skilled labor demand, m_s denotes commercial land demand, and we denote the location-sector-specific commercial land supply by $M_{\ell s}$. Entry cost technology differs across sectors through a sector-specific cost shifter τ_s and differences in factor shares η_s. For quantitative reasons, we include commercial land and exclude capital from the entry-cost function. As shown above, if entry costs increase with the cost of local factors, locations with higher prices for such factors have larger firms in equilibrium. However, the wage-density gradient in the data is not steep enough to generate the establishment-size-density gradient, suggesting local land must be part of the entry cost since its supply is less elastic than labor. For capital, the enormous decline in the capital price in the data would imply a dramatic drop in entry costs everywhere and nationwide decreases in average firm size. In contrast, the average firm size in the US data has increased moderately since 1980.

In line with much of the literature, we choose the distribution of firm heterogeneity to be Pareto with a scale parameter of 1 and a shape parameter ν so that:

$$\Omega_s(z) := 1 - z^{-\nu},$$

for $z > 1$, so that no differences in firm heterogeneity exist across sectors.[33]

3. Quantifying the Theory

We use the model as an accounting device to measure the variation in wages and employment across locations, sectors, and worker types due to the observed decline in the investment price of IT capital. In preparation, we estimate the model's structural parameters using a combination of model-implied estimating equations and indirect

[33] As a result, local differences in productivity shifters $\{A^h_{\ell s}, A^l_{\ell s}\}$ and factor prices drive all variation in firm scale and input choices across locations and sectors.

inference.

3.1 Calibrating the Model

For our calibration, we map locations in the model to the 722 commuting zones covering the entire continental US. We differentiate two sectors, Business Services and Other Sectors, which include all other private, non-agricultural employment. Following Krusell et al. (2000), we define "high-skill" workers as those with at least a four-year college degree and "low-skill" workers as all others; we refer to high-skill and low-skill workers as college- and non-college-educated workers from here on. We calibrate our model to an annual frequency.

Since we require the education status of workers, we use US Decennial Census data and American Community Survey data (see Ruggles, Genadek, Goeken, Grover, and Sobek, 2017) in our calibration. Using this source, we construct a panel of wages and employment for each commuting zone, sector, education group, and decade. Using the same data, we also create a residential rent-price index as the commuting-zone-year fixed effects in a regression of gross rents of individuals on a large set of dwelling characteristics.[34] We also obtain data on IT capital use by sector and prices of IT from the BEA. Since the Census data are decadal, we interpolate wages linearly to get an annual panel of local wages and employment across locations, sectors, and skill groups.

The model features time-varying *structural residuals*, $\{A^f_{\ell st}, A^k_s, B_{\ell ft}, B_{\ell sft}, O_{\ell t}, M_{\ell st}, \mathcal{Z}_t\}$, and constant *structural parameters*, $\{\bar{\epsilon}_s, \varphi_s, \sigma_s, \tau_s, \eta_s, \varrho^1_f, \varrho^2_f, \alpha_f, \iota_s, \nu, \zeta, \gamma, \beta, \xi\}$. Our estimation strategy does not assume the model is in a steady state between 1980 and 2015; we estimate most of the model's structural parameters by targeting moments along its out-of-steady-state path.[35] An advantage of our calibration strategy is that we can target cross-sectional moments between 1980 and 2015. Although most structural parameters are estimated jointly, we discuss each parameter's calibration strategy in terms of its most informative empirical moment.

An essential aspect of our estimation is that none of the structural parameters central to our mechanism is estimated to match *dynamic* moments. Given the structural parameters, we infer the structural residuals to match wages and employment by commuting zone, sector, skill type, year, and data on rents, capital prices, and capital value-added. Table 3 provides an overview of all calibrated parameters.

Non-homotheticity $\bar{\epsilon}_s$. In our theory, factor ratios vary with firm scale within location-sector pairs as long as the scale elasticities in equation (12) are non-zero. As a result, we estimate $\bar{\epsilon}_s$ to match the relationship between capital expenditure per worker and firm

[34] This procedure is standard in the literature, and we provide more information in the Online Appendix.
[35] After 2015, we hold all structural residuals fixed, so the economy eventually converges to a steady state.

size as measured by employment shown in Table 1. For each sector, we run a regression of log capital expenditure per worker on log firm employment within model and data in 2013 and chose $\bar{\epsilon}_s$ to ensure the coefficients on firm employment between model and data coincide. Note that our indirect inference procedure does not require unbiased estimates of the coefficients on firm size (Smith, 1993; Smith, 2008).

Recall that the ACES data combines capitalized and non-capitalized IT capital expenses. In the model, firms only make capitalized expenses since the capitalists make all (non-capitalized) investments. As a result, our estimation implicitly assumes that capitalized and non-capitalized expenses have a similar elasticity to firm size. The Spiceworks data supports this assumption.

The expressions for the scale elasticities in equation (12) show larger firms produce with higher capital-to-labor ratios than smaller firms as long as $\bar{\epsilon}_s < 0$ and $\sigma_s < 1$. To match the positive relationship between capital expenditure per worker and firm size in Table 1, our calibration finds that $\bar{\epsilon}_s = -0.21$ for Business Services and $\bar{\epsilon}_s = -0.08$ for other sectors. The more negative scale elasticity for Business Services reflects that the empirical relationship between capital expenditure per worker and firm size in Table 1 is stronger for the sector than for others.

Factor Substitution Elasticities σ_s and φ_s. Given the non-homotheticity parameter $\bar{\epsilon}_s$, the substitution elasticities σ_s and φ_s determine the ease of substituting between capital, college labor, and non-college labor at the firm level.

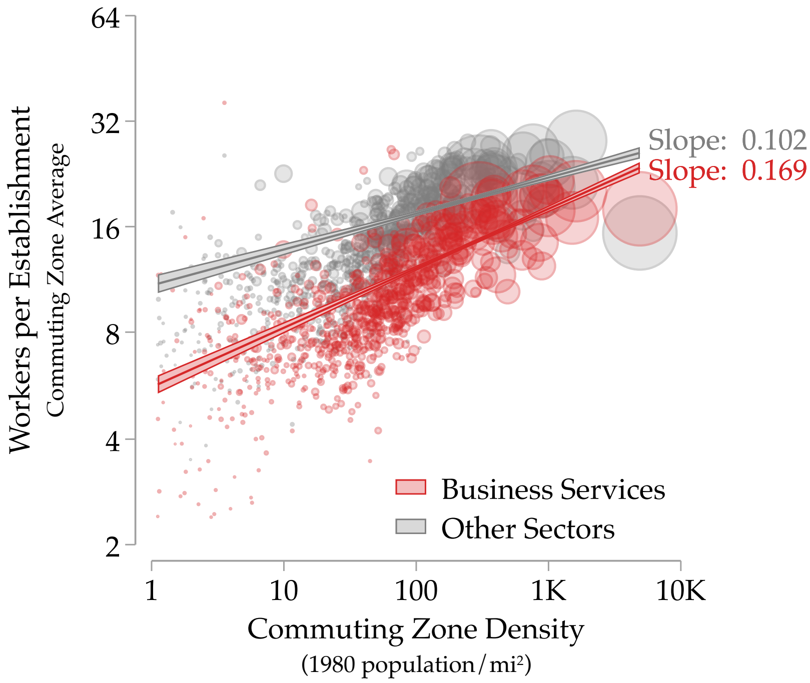
The scale elasticities in equation (12) show a non-zero non-homotheticity parameter $\bar{\epsilon}_s$ not only induces variation in the capital-labor ratio across firms of different sizes but also leads the ratio of high- to low-skill labor to vary with firm size. For a given non-homotheticity parameter $\bar{\epsilon}_s$, the substitution elasticities determine how the skill intensity varies with firm size. We exploit this in our estimation of the low-skill elasticity of substitution φ_s by choosing it to match how the ratio of high- to low-skill workers varies with firm size.

Using data from the 1992 Current Population Survey (CPS), we find positive elasticities of both sectors' high- to low-skill ratio to firm size in line with evidence from Trottner (2019).[36] For each sector, we regress the log of the high-to-low-skill labor ratio on log firm employment and choose φ_s to ensure that the model matches coefficients in the same regression in the data. We find that $\varphi_s = 1.29$ for Business Services and $\varphi_s = 1.52$ for other sectors so that low-skill labor is moderately substitutable with the high-skill-capital bundle in both sectors.

Given φ_s and $\bar{\epsilon}_s$, the elasticity of substitution σ_s pins down the *aggregate* elasticity of substitution between IT capital and all types of labor. The model does not deliver

[36] Figure OA.14 in the Online Appendix. Trottner (2019) also documents similar skill-intensity to firm-size elasticities in German microdata.

FIGURE 8: AVERAGE ESTABLISHMENT SIZE ACROSS COMMUTING ZONES, 1980



Notes: The figure shows the average number of workers per establishment within each commuting zone (Tolbert and Sizer, 1996) and sector. The slope numbers indicate the coefficient on log commuting zone density in an employment-weighted regression of log average establishment size on the log commuting zone population density; the line shows the fitted regression lines with 95% confidence intervals. Circle size is proportional to the commuting zone population. The underlying data comes from the Quarterly Census of Employment and Wages published by the US Bureau of Labor Statistics.

a closed-form expression for the capital-labor elasticity on the firm level or in the aggregate. We let σ_s be constant across sectors and calibrate it to match the aggregate elasticity of substitution between IT capital and labor of 0.95 estimated in Lashkari et al. (2024). We find that $\sigma = 0.49$, making IT capital and high-skill labor strong complements on the firm level.

Oberfield and Raval (2021) estimate an aggregate elasticity of substitution between capital and labor in manufacturing between 0.5 and 0.7. An alternative estimate comes from Karabarbounis and Neiman (2014), putting the elasticity at 1.25 across all sectors of the economy. For our counterfactual analysis below, we present robustness checks showing how our analysis changes when we target each of these alternative macro elasticities instead.

Entry cost parameters, η_s and τ_s. The free-entry condition implies firms in locations with higher entry costs must make greater variable profits and, hence, operate at a larger scale. As a result, the level and cross-location variation of entry costs determine aggregate and cross-location firm-size patterns in our model. Figure 8 shows the average number of workers per establishment across commuting zones in 1980, separately for Business Services and other sectors. The graph uses public data from the Quarterly Census of Employment and Wages (QCEW).

We choose the level of entry costs in each sector, τ_s, to match the average establishment size in each sector in the aggregate economy in 1980. We choose the labor share in entry costs, η_s, to match the gradient of average firm size with respect to population density separately for each sector. Commercial land prices increase faster with population density than labor since, unlike labor, land supply is inelastic on the local level. As a result, the higher the expenditure share on commercial land in entry costs $(1 - 2\eta_s)$, the steeper the cross-sectional relationship between firm size and density in a sector.

We find $\eta_s = 0.08$ for Business Services and $\eta_s = 0.05$ for the other sectors, showing that simply allowing wages to determine entry costs is insufficient to produce the establishment-size gradient seen in the data. The gradient of wages with respect to population density of about 0.04 in 1980 Census data is not steep enough to generate sufficient equilibrium variation in profits and, hence, firm size.[37] The fact that the land share is larger in Business Services reflects that the sector's establishment-size to population density gradient is much steeper in the data (cf. Figure 8).

Labor-supply elasticities, ϱ_f^1 and ϱ_f^2. We estimate sectoral and spatial labor-supply elasticities for workers using structural estimating equations implied by the model.

We begin by estimating the sectoral labor-supply elasticities using variation across sectors within commuting zones. Taking logs of equation (7) and writing it in changes yields the following estimating equation:

$$\Delta \log \mu_{\ell sft} = \varrho_f^2 \Delta \log w_{\ell sft} + \Delta \log \left(\sum_{s'} B_{\ell s' ft} w_{\ell s' ft}^{\varrho_f^2} \right) + \Delta \log B_{\ell sft}, \tag{13}$$

which we can estimate "outside" the model to recover the model-consistent labor-supply elasticity. Equation (13) shows the coefficient on the wage-change term identifies the labor-supply elasticity. The second term is a location-type-time fixed effect. The third term is a structural residual that highlights the need to instrument for wage changes to recover ϱ_f^2: if wage growth correlates with changes in the unobserved sectoral amenities in a location, an ordinary least squares (OLS) regression yields biased estimates of ϱ_f^2.

We construct a Bartik-like "predicted" wage change for each location-sector-type triplet and use it as an instrumental variable (IV). For each location-sector-type and period, we compute the wage growth for that sector type outside the location during the same period, leaving out the sectoral wage and employment of the location-sector pair. The exclusion restriction requires that, controlling for commuting zone-year effects, average wage changes outside a commuting zone are uncorrelated with changes in type-specific amenities within a commuting zone.

[37] Recall that the model matches the full panel of wages exactly so that the wage-density gradient from the data is the same as the wage-density gradient in the model.

TABLE 2: SECTORAL AND SPATIAL LABOR-SUPPLY ELASTICITIES

	(1) College	(2) Non-College	(3) College	(4) Non-College
PANEL A: SECTORAL ELASTICITIES				
Δ log(Wage)	0.0960	1.305***	0.691**	0.444***
	(0.126)	(0.0777)	(0.215)	(0.118)
N	17713	17782	17713	17782
First Stage F	108.1	640.9	146.8	360.6
Fixed Effect: Year-Commuting-Zone	✓	✓	✓	✓
Commuting Zone Pop. Weighted			✓	✓
PANEL B: SPATIAL ELASTICITIES				
Δ log(Deflated Wage Index)	4.796***	3.579*		
	(1.065)	(1.780)		
Δ log(Wage Index)			4.100***	3.023***
			(0.889)	(0.545)
N	2223	2223	516	516
First Stage F	20.27	4.042	28.02	37.37
Fixed Effect: Year	✓	✓	✓	✓
Instrumented Rent			✓	✓

Notes: * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. Robust standard errors in parentheses. This table implements the structural labor supply equations for sectors and locations in the data. The underlying data comes from the US Decennial Census Files and, after 2000, from the American Community Survey. We run regressions in decadal differences and instruments for wage changes using the instrumental variables described in the body of the paper. Panel A shows the regressions for the sectoral labor supply elasticities based on equation (13). Columns 1 and 2 show coefficient estimates from a regression that uses data on all NAICS-1 sectors. Columns 3 and 4 show the same regressions but weights by commuting zone population in 1980. Panel B shows the regressions for the spatial labor supply elasticities based on equation(14). Columns 1 and 2 show coefficient estimates from a regression that uses our full sample. Columns 3 and 4 show estimates that follow the two-instrument procedure inDiamond (2016) and use a smaller sample for which the requisite data is available.

Estimating equation (13) with location-type-time fixed effects means we have to rely on cross-sector variation within each location and education group. To increase the statistical power of our estimation, we estimate equation (13) across NAICS-1 sectors instead of grouping all non-Business Services sectors into a single other category.[38]

Panel A of Table 2 presents the resulting elasticity estimates. Columns 1 and 2 present the estimated coefficients on the wage-change term in unweighted regressions, whereas Columns 3 and 4 weight each regression by commuting zone population. Columns 3 and 4 are our preferred estimates, and they suggest slightly higher sectoral elasticities for college-educated workers than those without a college degree. Our elasticities are at the high end of the 0.2-0.7 range implied across different specifications in Artuç, Chaudhuri, and McLaren (2010). However, those authors pool their estimates over

[38]Our identification strategy is model-consistent because if a model version with many sectors generated the data, we can identify the true labor-supply elasticity even when aggregating across some sectors. The intuitive reason is that the labor-supply elasticity of any individual sector only depends on the parameter ϱ_f^2 and the local employment share of that sector, not the individual employment shares of all other sectors.

worker types.

Similar to the sectoral elasticities, the model also implies an estimating equation for the spatial labor-supply elasticity. In particular, taking logs of the first equation in expression (7) and differencing across time, we obtain

$$(14) \qquad \Delta \log \lambda_{\ell f t} = \varrho_f^1 \Delta \log\left(r_{\ell t}^{-\alpha_f} \Psi_{\ell f t}\right) - \Delta \log \sum_{\ell} B_{\ell f t} (r_{\ell t}^{-\alpha_f} \Psi_{\ell f t})^{\varrho_f^1} + \Delta \log B_{\ell f t}.$$

The spatial labor-supply elasticity appears as the coefficient of the change in the wage index $\Psi_{\ell f t}$ deflated by the rental rate. The second term on the right is a type-time fixed effect. The third term is the change in the location-type-specific amenity. Since these amenities are unobserved, estimating equation (14) with OLS may yield biased estimates of ϱ_f^1. As a result, we construct an instrument for the first term in the equation.

We can construct the deflated wage index for each location-type-time using previously estimated parameters and data. In particular, given our estimate of the sectoral supply elasticity, observed wages, and the implied amenity residuals, we can construct $\Psi_{\ell f} = (\sum_s B_{\ell s f} (w_{\ell s f}^L)^{\varrho_f^2})^{1/\varrho_f^2}$ (cf. equation (7)). In addition, we estimate α_f directly from Decennial Census data by dividing mean annual rental payments by mean annual income for each commuting zone. The term $r_{\ell t}$ corresponds to our commuting zone rental price index.

We instrument for changes in the deflated index using another Bartik-like IV constructed as follows:

$$\text{IV}_{\ell f} = \sum_s \mu_{\ell s f} \Delta w_{sf,-\ell}^L,$$

where $w_{sf,-\ell}$ the average sectoral wage for type-f workers in all locations except location ℓ itself. The exclusion restriction requires that initial employment shares are uncorrelated with changes in local amenities between two time periods.[39]

Columns 1 and 2 of Panel B of Table 2 present the result of this IV regression using 10-year differences from 1980-2010 over all 722 commuting zones.

Because the instrument has a weak first stage for non-college-educated workers, we additionally follow the estimation strategy in Diamond (2016) in Columns 3 and 4. Diamond (2016) interacts a similar Bartik-type IV with exogenous land-use regulation and available developable land from Saiz (2010) to construct a second, separate instrument for the rent term. The estimation requires data on land-use regulation from Saiz (2010), which is only available for select Metropolitan Statistical Areas (MSA) and not for our commuting zones. We adopt the estimation procedure in Diamond (2016) to our context

[39] Goldsmith-Pinkham, Sorkin, and Swift (2020) discuss how to estimate spatial labor-supply elasticities using US Census data. Our setup deviates from theirs through an explicit control for rent and since we run separate regression for each worker type. We implement an alternative approach more akin to the setup in Goldsmith-Pinkham et al. (2020) in the Supplementary Material and find similar results.

and estimate the elasticities on the MSA level using the two-instrument approach.

Columns 3 and 4 Panel B of Table 2 present the results. Since there are fewer MSAs than commuting zones, the sample size in Columns 3 and 4 is smaller than in Columns 1 and 2.[40] Reassuringly, our results in Columns 1 and 2 and Columns 3 and 4 are similar, and we use the first two columns in our baseline.

Note that the labor-supply elasticities are the only structural parameter we estimate using data over time. Since our mechanism operates through the labor demand side, the labor-supply elasticities only affect its propagation through quantities or prices, but not its strength per se. Recall from our theory that the free-entry condition determines average wages in each location and sector. As a result, the labor-supply elasticities thus only determine the relative wage response of high- versus low-skill workers.[41]

Other Structural Parameters. We set $\iota_s = 4$, in line with consumer goods from Hottman, Redding, and Weinstein (2016), and similar to values used for the LBD in Garcia-Macia, Hsieh, and Klenow (2019). We assume varieties have the same elasticity of substitution within and across sectors so that $\gamma \to \infty$. We use Census data from 1980-2015 for the average rent payments to income to set the Cobb-Douglas share for housing α_f to 0.18 for college-educated workers and 0.32 for non-college-educated workers. We choose the tail coefficient of the firm-efficiency distribution, ν, to match the tail coefficient of the establishment size distribution in the LBD, which we estimate to be 1.2, consistent with Cao, Hyatt, Mukoyama, and Sager (2017). We set the capitalist discount rate β to 0.97 to match a long-run interest rate of 3%. Finally, we set the firm exit rate $\xi = 0.1$ to match the exit rate in the LBD in 1980.

Productivities, Amenities, and Land Supply. Given all other structural parameters, we infer the structural residuals of our model to match a variety of data moments exactly (see Redding and Rossi-Hansberg, 2017). In particular, we infer productivities $(A^h_{\ell st}, A^l_{\ell st})$ and amenities $(B_{\ell ft}, B_{\ell sft})$ to ensure the model matches average wages and employment counts for all locations, sectors, and worker types exactly every year.

We choose the productivity of IT capital in each sector, A^k_s, to match the value-added share of IT capital in each sector which we compute using data from the BEA. We infer the productivity of IT capital production, $\mathcal{Z}_t$, to match the time series of the price of IT capital from the BEA asset tables. Recall that in our model, p^K_t is the investment price of capital, which we can directly observe in the data.

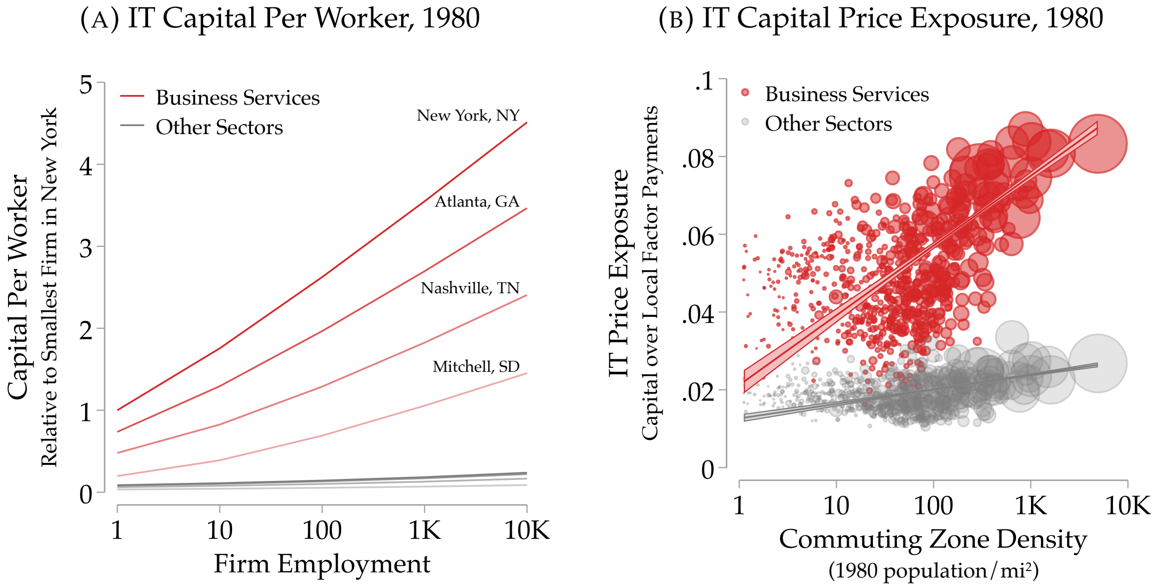
We choose the inelastic residential housing supply in each region, $O_{\ell t}$, so that the model matches our commuting-zone-year price index exactly.[42] Because we do not

[40] Note that our estimates deviate slightly from those in Diamond (2016) since their estimation features additional restrictions in a Generalized Method of Moments estimator.

[41] Having separate labor-supply elasticities by skill group is also not crucial, and in the Online Appendix we show the robustness of our main model exercise to assuming the identical elasticities across groups.

[42] Adding endogenous housing supply is trivial and isomorphic to our current setup with a different

FIGURE 9: IT CAPITAL PER WORKER AND IT PRICE EXPOSURE IN THE MODEL



Notes: The figure presents the output from our calibrated model for 1980. The left panel plots the policy functions that map firm size to optimal capital per worker within a set of illustrative locations. Factor prices are constant across firms within a location-sector. We normalize capital per worker to 1 for firms with one employee in Business Services in New York. The right panel plots the exposure statistic to IT price changes for each location-sector pair as implied by Theorem 1. The exposure statistic equals the ratio of total payments to IT capital relative to total payments to labor and commercial land.

observe commercial land prices for the cross-section of commuting zones, we assume the commercial land supply in each location-sector, $M_{\ell st}$, is proportional to that location's residential land supply. The constant of proportionality and entry-cost shifter, τ_s, are not separately identified; we set the constant to 1 without loss of generality.

3.2 Location-Sector Exposure in the Calibrated Model

Using the calibrated model, we construct the policy functions that map firm size into capital per worker within each location. The left panel of Figure 9 plots these policy functions for a handful of representative cities for the Business Services sector. The model predicts that firms in higher-density locations like New York spend more on capital per worker than firms of the same size elsewhere.

Next, we use the calibrated model to construct the exposure statistic implied by Theorem 1 in 1980. The right panel of Figure 9 shows the ratio of total payments to IT capital relative to total payments to labor and commercial land by sector and location in 1980. Theorem 1 shows this ratio is a sufficient statistic for the exposure of local factor prices in a location-sector to changes in IT prices.[43]

Exposure to IT price changes strongly increases with population density for the Business

labor-supply elasticity.

[43]Note that since our model has one type of capital, the ratio summarizes the location-sector exposure to direct and indirect effects.

TABLE 3: OVERVIEW OF MODEL PARAMETERIZATION

	Estimated Structural Parameters	*Value*	*Description of Moment*	*Moment: Model/Data*
$\bar{\epsilon}_s$	Non-homotheticity	(-0.21, -0.08)	2013 Elasticity of capital per worker to firm size	(0.29, 0.12)/ (0.29, 0.12)
φ_s	EoS Skilled and Unskilled Labor	(1.29, 1.52)	1992 Elasticity of high-low skill ratio to firm size	(0.05, 0.09)/(0.05, 0.09)
σ	EoS Skilled Labor and IT Capital	0.49	2007 Macro-Elasticity from Lashkari et al. (2024)	0.95/0.95
τ_s	Entry-Cost Shifter	(5.3, 843)	1980 Average establishment size by sector	(11.7, 14.4)/(11.7, 14.4)
η_s	Labor Share in Entry Cost	(0.08, 0.05)	1980 Elasticity of estab. size to pop. density	(0.17, 0.10)/(0.17, 0.10)
ϱ_f^1	Spatial Labor-Supply Elasticity	(3.6, 4.8)	Estimated using equation (14)	N/A
ϱ_f^2	Sectoral Labor-Supply Elasticity	(0.44, 0.69)	Estimated using equation (13)	N/A
α_f	Housing share in final consumption	(0.32, 0.18)	1980-2015 Avg. rent payments over income	(0.32, 0.18)/(0.32, 0.18)
ν	Pareto Shape Parameter	7.3	Tail parameter of LBD estab. size dist.	(1.2)/(1.2)
ξ	Firm Exit Rate	0.1	1980 LBD Exit Rate	(10%)/(10%)
	External Structural Parameters	*Value*	*Source*	*Moment: Model/Data*
ι_s	Sectoral Demand Elasticity	(4, 4)	Hottman et al. (2016)	N/A
γ	Sectoral Elasticity of Substitution	∞	N/A	N/A
β	Capitalist Discount Rate	0.97	Interest rate of 3%	N/A
	Productivities, Amenities, and Land	*Value*	*Data Matched*	*Moment: Model/Data*
$A_{\ell st}^f$	Location Productivity Shifter	Various	1980-2015 employment and wages	Various
A_s^k	Sectoral IT Capital Productivity	Various	1980 IT share of value added by sector	Various
$B_{\ell ft}$	Location Amenities	Various	1980-2015 Commuting Zone employment and wages	Various
$B_{\ell sft}$	Sectoral Amenities	Various	1980-2015 Commuting Zone employment and wages	Various
$O_{\ell t}$	Residential Land Supply	Various	1980-2015 local residential rent index	Various
$M_{\ell st}$	Industrial Land Supply	Various	Proportional to residential land supply	Various
$\mathcal{Z}_t$	Productivity of IT production	Various	1980-2015 BEA IT capital price index	Various

Notes: This table shows the baseline parameterization of the model. The productivity, amenity, and land supply terms vary across locations, sectors, and factor types, so their values are not listed. Where two values appear for a sector-specific parameter, the value for Business Services is first. Where two values appear for an education-group-specific parameter, the value for low-skill appears first.

Services sector and is essentially flat for other sectors. The level difference in exposure across sectors reflects the different productivity of IT capital in two sectors in 1980 (A^k_s). We infer a much lower productivity of IT capital outside Business Services since our targeted moment, capital payments relative to labor payments in 1980, is much lower for other sectors than the Business Services sector.

How exposure increases with population density in each sector reflects the net effect of the horserace between the neoclassical channel and the scale channel outlined in Section 2.3.[44] Since we estimate that capital and high-skill labor are strong complements, the neoclassical channel is strong in both sectors and pushes for lower exposure in high-wage, high-density commuting zones. An even stronger scale channel counteracts the neoclassical channel for Business Services and, to some extent, for other sectors. The scale channel is strong in Business Services because we estimated a strong non-homotheticity and an entry cost that increases rapidly with density.

The exposure elasticity in the right panel of Figure 9 combined with Theorem 1 suggests a decline in the price of IT capital should have a strongly urban-biased effect on local factor prices in the Business Services sector and much less so for other sectors. However, the theorem only applies for small changes in investment prices. To understand the effect of the large observed decline in the investment price of IT capital between 1980 and 2015 on wages across locations and sectors, we solve the full dynamic path of our model in general equilibrium.

4. Accounting for Urban-Biased Growth

We now quantify how much of the observed urban-biased growth in the US economy is accounted for by the decline in the investment price of IT capital in the data.

Our calibration exercise produces a set of preference and technology structural residuals to account for the full panel of wages and employment counts across commuting zones, sectors, and worker types between 1980 and 2015. Our accounting exercise isolates the changes in wages and employment across commuting zones, sectors, and worker types due to the observed decline in IT prices alone.

In particular, we compute the counterfactual dynamic path of an "IT-only" economy in which only the productivity of capital production ($\mathcal{Z}_t$) varies, which we inferred to match the time series of IT prices; all other preference and technology structural residuals remain at their 1980 levels. Changing the productivity of capital production ($\mathcal{Z}_t$) from its 1980 to 2015 value implies the price of IT capital in our counterfactual

[44] Relative to the simple model, the quantitative model has two types of labor: one complementary with capital and the other substituting. High wages decrease exposure for the complementary type of labor; for the substitutable type of labor, the opposite is the case. The net effect depends on the relative cost shares of the two types of labor.

TABLE 4: WAGE-DENSITY ELASTICITY IN DATA AND MODEL

	Data		IT-Only Economy 2015			
	1980	2015	Base	$\bar{\epsilon}_s = 0$	End. A	End. B
Business Services	0.070	0.154	0.151	0.072	0.151	0.150
Other Sectors	0.060	0.070	0.069	0.060	0.069	0.068
Aggregate	0.063	0.102	0.103	0.067	0.103	0.101
Δ Aggregate		0.039	0.039	0.003	0.039	0.039

Notes: This table shows the regressions of log average wages on population density in the cross-section of US commuting zones in the data and in the IT-Only economy. Note that the 1980 cross-section is the same in the data and the IT-Only economy. The data underlying the data columns comes from the 1980 Decennial Census and the 2015 American Community Survey. Data Column 3 shows the wage-density elasticity in the baseline IT-only economy. Data Column 4 shows the wage-density elasticity in a homothetic version of the baseline IT-only economy with $\bar{\epsilon}_s = 0$, other structural parameters the same, but regional fundamentals re-calibrated. Data Column 4 shows the wage-density elasticity when local productivity terms are increasing functions of local population density. Data Column 5 shows the wage-density elasticity when local amenity terms for college workers are increasing functions of the local college share of employment, as in Diamond (2016).

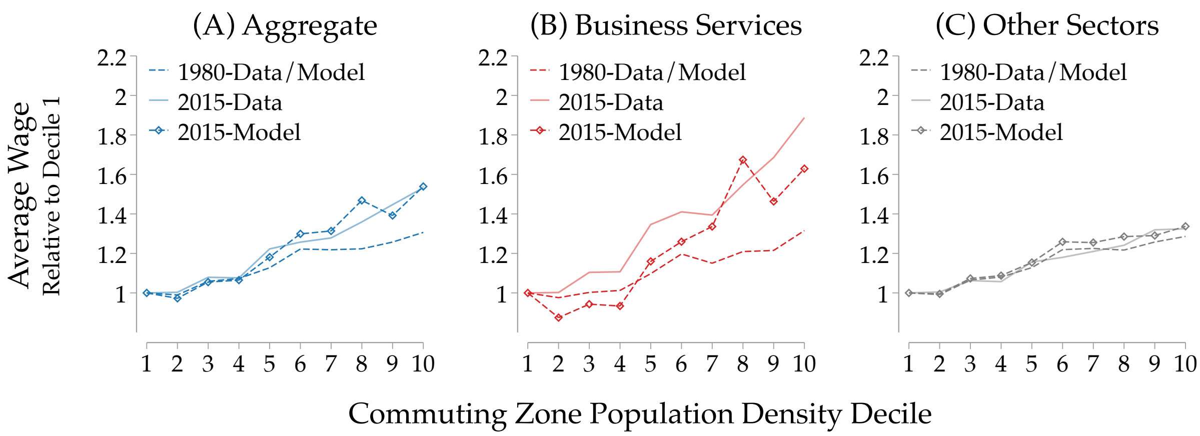
economy replicates the path of the price in the data. In addition, we adjust the *aggregate* college employment share as in the data to enable us to directly compare employment shares across locations and sectors between our counterfactual economy and the data. We emphasize that the IT-only economy differs from the 1980 economy only through changes in two numbers: the aggregate college share and the productivity of capital production.

The first two columns of Table 4 present the wage-density gradient in the data in 1980 and 2015, both in the aggregate and separately for each sector. Column 3 shows the counterfactual wage-density gradient in 2015 in our baseline "IT-only" economy. Declining IT prices alone can explain almost all of the increase in the wage-density gradient. The increase in the gradient occurs primarily in the Business Services sector, exactly as in the data.

The decline in IT prices has different effects across sectors because of the exposure differences across sectors shown in Figure 9. Figure 10 replicates Figure 1 using data from our "IT-only" economy. Note that because we have to use Decennial Census data for the calibration, the wage-density gradients in the data in both years differ somewhat from Figure 1.

The increase in the aggregate gradient reflects the changes in the gradient in each sector and the reallocation of employment across sectors. The left panel of Figure 11 shows the college share of employment in each sector in the model and the data. The right panel of Figure 11 shows the college share of employment across commuting zones and sectors. The "IT-only" economy generates almost the full degree of reallocation in the actual data. It produces a slightly flatter gradient in the ratio across locations,

FIGURE 10: URBAN-BIASED GROWTH ACROSS IN 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.

particularly in the highest-density deciles, including the cities of New York and Chicago. Note that although we adjust the *aggregate* share of college-educated workers as in the data, all sorting of workers across sectors and locations is an endogenous response to the changes in the price of IT capital. Lastly, because the model replicates the changes in worker stocks across locations and reproduces the change in the wage gradients, it also generates much of the urban bias in the changes in the residential rent-price index observed in the data.[45]

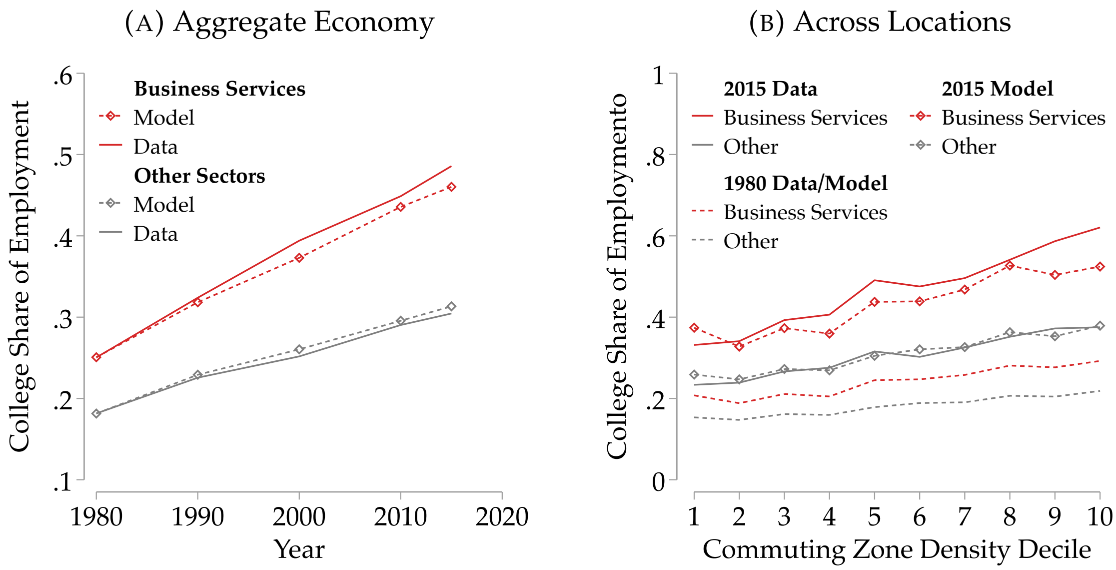
To further illuminate the mechanism translating aggregate changes in IT prices into unbalanced wage growth across regions, we compare an economy with and without non-homotheticity in production. Our theory showed the non-homotheticity of the production function is essential in generating urban-biased growth. The non-homotheticity gives rise to the scale channel at the heart of big cities' exposure to IT price declines (cf. Figure 9). In Column 4 of Table 4, we present the wage-density gradient in 2015 in a homothetic version of the model, with the non-homotheticity parameter set to zero ($\bar{\epsilon}_s = 0$), and re-calibrated technology and preference residuals. Without the non-homotheticity, the wage-density gradient does not meaningfully increase as IT prices fall.[46]

Figure 11 shows that the model captures the split into quantity and price responses caused by the IT price decline. Figure 13 additionally replicates the decomposition from Fact 1 in Section 1 in the model-generated data. The vast majority of urban-biased growth that the IT price decline generates comes from initial differences in Business

[45]See Figure OA.15 in the Online Appendix.

[46]Shutting down the non-homotheticity does not entirely eliminate the scale channel. The second component of the scale channel remains active since entry costs do not use capital, explaining why wage growth is flat rather than rural-biased when the non-homotheticity parameter is set to zero.

FIGURE 11: SKILL DEEPENING IN MODEL AND DATA



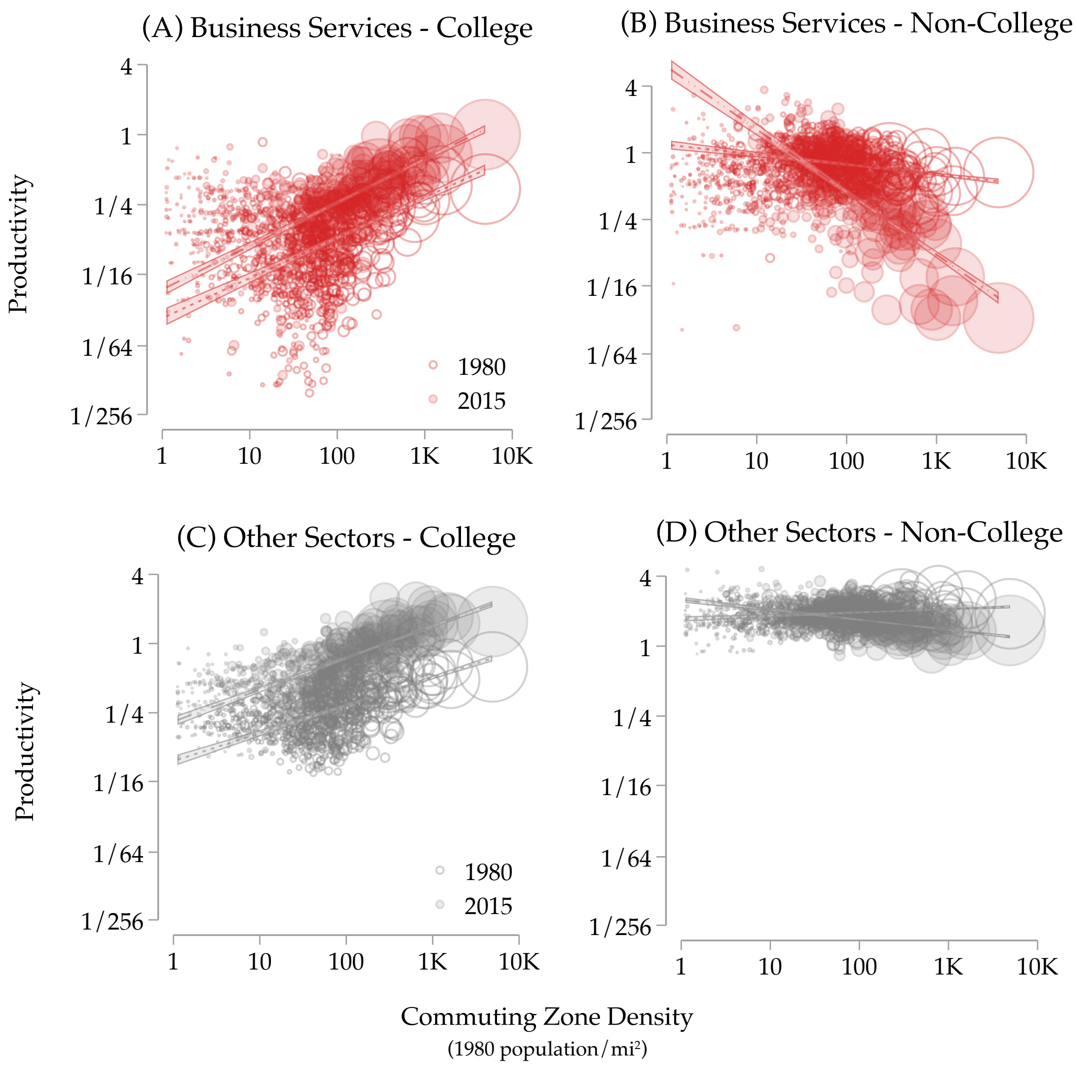
Notes: The left panel of this figure shows the growth in the ratio of college-educated to non-college-educated workers in both the model and the Decennial Census data by year and sector. The right panel of this figure shows this ratio in 2015 in both the model and the data by sector across the commuting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

Services employment shares interacting with faster wage growth in the sector in higher-density locations. The share accounted for by such wage-growth differences is 39% in the model and 44% in the data.

In the Online Appendix, we also replicate the decomposition of urban-biased growth into the contributions of large and small firms shown in Figure 4. Although the model captures firm-size differences in total payroll growth well, it generates too much from employment growth and too little from wage growth. Recall that all firms pay the same wage within location and sector, even though large firms pay higher wages for the same type of worker in reality. Allowing for firm-specific labor-supply curves as in Trottner (2019) is a straightforward extension that could help the model speak to these differences.

While the change in the price of IT capital can generate most of the observed urban-biased growth, the structural technology residuals are important for much of the observed *aggregate* wage growth. Figure 12 shows the technology structural residuals in 1980 and 2015 for each sector and education group. For college-educated workers in all sectors, aggregate productivity has increased and represents a key source of aggregate wage growth. However, the structural residuals are not a source of urban-biased growth in the Business Services sector. In fact, the residuals show rural-biased productivity growth for low-skill workers in both sectors. These patterns partially reflect sub-industries of the Business Services sector, such as business support services (NAICS 56) that predominantly use low-skill labor and increasingly set up establishments in

Figure 12: Structural Productivity Residuals by Sector and Skill



Notes: The paper shows the calibrated productivity residuals for each sector and skill group in 1980 and 2015. The x-axis shows workers per square mile. We scale all productivity terms by the productivity of college-educated workers in the Business Services sector in the New York commuting zone in 2015.

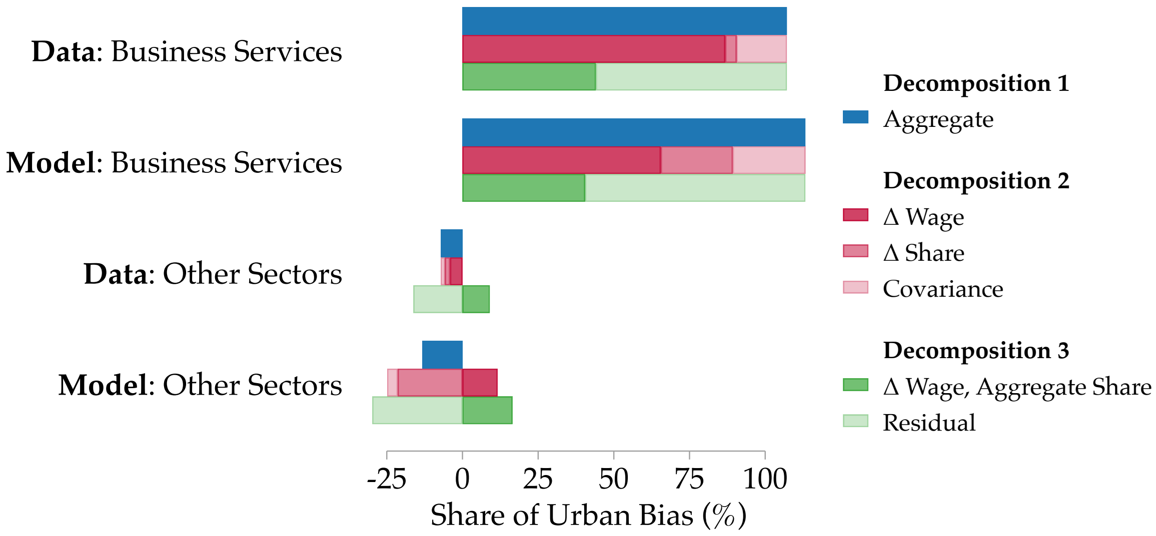
low-density locations, as described in Liao (2012). Outside the Business Services sector, it highlights the manufacturing sector's decline, which happened most rapidly in big cities (see Autor, 2019) and impacted low-skill workers most.

In the Online Appendix, we compute additional counterfactual economies, in which we vary only the technology residuals or the amenity residuals. Confirming the patterns in Figure 12, we find that changes in neither set of residuals generate any urban bias in wage growth.[47] The decline in IT prices is the sole force in the economy generating meaningfully urban-biased wage growth.

An extensive literature in urban economics has emphasized the endogenous nature of local productivity and amenities. Our calibration procedure infers technology and preference residuals to match annual data and is agnostic about their endogenous

[47]See Table OA.3 in the Online Appendix.

FIGURE 13: DECOMPOSING URBAN-BIASED GROWTH IN MODEL AND DATA



Notes: We use the decompositions from Section 1 for urban-biased growth in the model-generated counterfactual where only the price of IT and the aggregate worker stocks by skill adjust as in the data. The top panel shows the decomposition in the data; the bottom panel shows the same decomposition in the model-generated data.

nature. That is, it is completely consistent with a world where urban agglomeration economies and amenities that respond to skill supply are important features of reality.

However, our growth accounting exercise may change when amenity and productivity terms are partly endogenous. To understand how it would change, we repeat our accounting "IT-only" counterfactual in the presence of endogenous productivity and amenity terms. We use estimates from the meta-study of Ahlfeldt and Pietrostefani (2019) in allowing the labor productivity shifters $\{A^h_{\ell s} A^l_{\ell s}\}$ to increase in total local employment, a classic "agglomeration" spillover. In addition, we follow Diamond (2016) and model positive spillovers in amenities $B_{\ell f}$ for college-educated workers from the presence of other college-educated workers, using her estimate for the spillover parameter.[48]

Column 5 of Table 4 presents the resulting wage-density gradients in 2015 with productivity spillovers. The result is virtually identical because the spillover parameter in the meta-study of Ahlfeldt and Pietrostefani (2019) is small and total city populations change little in our counterfactual.

In response to the decline in IT prices, the college share among workers in high-density locations increases. With endogenous amenities as in Diamond (2016), these inflows lead to higher amenities and, hence, lower high-skill wages in spatial equilibrium, offsetting some of the urban-biased wage growth that would otherwise occur. Column 6 shows the wage-density gradient generated by the decline in IT prices in the version of the model with endogenous amenities. The model predicts a slightly smaller increase in the

[48]The Online Appendix describes provides additional details.

wage-density gradients once productivities endogenously adjust, because workers do not need as much monetary compensation to live in big cities when amenities improve.

Finally, Table OA.3 in the Online Appendix provides additional robustness checks for our accounting exercise. We show how the contribution of the change in IT prices to urban-biased growth depends on the value for the aggregate elasticity of substitution between capital and labor we target in our calibration of σ. When we target an elasticity of 0.65 in Oberfield and Raval (2021), the decline in IT prices generates only about half of the urban-biased growth seen in the data, when we target 1.25 from Karabarbounis and Neiman (2014) it explains more than three times the observed urban-biased growth.[49] A lower elasticity of substitution means a stronger neoclassical channel, which lowers exposure elasticity in high-density locations; a higher elasticity weakens the neoclassical channel or even makes it contribute to the higher exposure of high-density locations as outlined in Section 2.3. Table OA.3 also shows urban-biased growth in the IT-only economy when college- and non-college-educated workers have the same labor-supply elasticities (set to the average values across groups). Our findings are virtually unchanged, suggesting heterogeneous labor-supply elasticities are not essential for understanding urban-biased growth.

Conclusion

Recent economic growth has been strikingly biased toward the richest and largest cities in the US. This paper shows that understanding why requires focusing on large establishments in skill- and information-intensive Business Services industries. These service firms have been key beneficiaries of innovation in information technology, which they used to scale up operations in the most productive US cities. A better understanding of these services can unlock new perspectives on the nature of economic growth in knowledge economies, and the accompanying inequality between workers and locations.

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. Domínguez-Iino (2022): "Location Sorting and Endogenous Amenities: Evidence from Amsterdam," *University of Chicago, Becker Friedman Institute for Economics Working Paper*.

[49]Note that both Oberfield and Raval (2021) and Karabarbounis and Neiman (2014) provide aggregate elasticities for all types of capital, not IT capital specifically. Our baseline calibration targeted an elasticity specific to IT capital from Lashkari et al. (2024).

ANDERSON, J. E., M. LARCH, AND Y. V. YOTOV (2020): "Transitional Growth and Trade With Frictions: A Structural Estimation Framework," *The Economic Journal*, 130, 1583–1607.

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. (2019): "Work of the Past, Work of the Future," Working Paper 25588, National Bureau of Economic Research.

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.

BAUM-SNOW, N. AND R. PAVAN (2013): "Inequality and City Size," *Review of Economics and Statistics*, 95, 1535–1548.

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.

BILAL, A. AND E. ROSSI-HANSBERG (2023): "Anticipating Climate Change Across the United States," Tech. rep., National Bureau of Economic Research.

BLOOM, N., M. DRACA, AND J. VAN REENEN (2016): "Trade Induced Technical Change? The Impact of Chinese Imports on Innovation, IT and Productivity," *The Review of Economic Studies*, 83, 87–117.

BRESNAHAN, T. F., E. BRYNJOLFSSON, AND L. M. HITT (2002): "Information Technology, Wrkplace Organization, and the Demand for Skilled Labor: Firm-level Evidence," *The Quarterly Journal of Economics*, 117, 339–376.

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

CAO, D., H. R. HYATT, T. MUKOYAMA, AND E. SAGER (2017): "Firm Growth through New Establishments," *Available at SSRN 3361451*.

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., D. LASHKARI, AND M. MESTIERI (2021): "Structural Change with Long-run Income and Price Effects," *Econometrica*, 89, 311–374.

COUTURE, V. AND J. HANDBURY (2020): "Urban Revival in America," *Journal of Urban Economics*, 119, 103267.

DAVIS, D. R. AND J. I. DINGEL (2020): "The Comparative Advantage of Cities," *Journal of International Economics*, 123, 103291.

DAVIS, D. R., E. MENGUS, AND T. K. MICHALSKI (2020): "Labor Market Polarization and The Great Divergence: Theory and Evidence," Tech. rep., National Bureau of Economic Research.

DESMET, K., D. K. NAGY, AND E. ROSSI-HANSBERG (2018): "The Geography of Development," *Journal of Political Economy*, 126, 903–983.

DESMET, K. AND E. ROSSI-HANSBERG (2014): "Spatial Development," *American Economic Review*, 104, 1211–1243.

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

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

EECKHOUT, J., C. HEDTRICH, AND R. PINHEIRO (2021): "IT and Urban Polarization," *CEPR Discussion Paper No. DP16540*.

EHRLICH, M. V. AND H. G. OVERMAN (2020): "Place-based Policies and Spatial Disparities across European Cities," *Journal of Economic Perspectives*, 34, 128–49.

FOGLI, A., V. GUERRIERI, M. PONDER, AND M. PRATO (2023): "The End of the American Dream? Inequality and Segregation in US Cities," Tech. rep., National Bureau of Economic Research.

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

GANAPATI, S. (2024): "The Modern Wholesaler: Global Sourcing, Domestic Distribution, and Scale Economies," *American Economic Journal: Microeconomics*.

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.

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

GOLDSMITH-PINKHAM, P., I. SORKIN, AND H. SWIFT (2020): "Bartik Instruments: What, When, Why, and How," *American Economic Review*, 110, 2586–2624.

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

HOTTMAN, C. J., S. J. REDDING, AND D. E. WEINSTEIN (2016): "Quantifying the Sources of Firm Heterogeneity," *The Quarterly Journal of Economics*, 131, 1291–1364.

JAIMOVICH, N. AND H. E. SIU (2020): "Job Polarization and Jobless Recoveries," *Review of Economics and Statistics*, 102, 129–147.

JIANG, X. (2023): "Information and Communication Technology and Firm Geographic Expansion," *CESifo Working Paper*.

KARABARBOUNIS, L. AND B. NEIMAN (2014): "The Global Decline of the Labor Share," *The Quarterly Journal of Economics*, 129, 61–103.

KLEINMAN, B. (2022): "Wage Inequality and the Spatial Expansion of Firms," Tech. rep., mimeo.

KLEINMAN, B., E. LIU, AND S. J. REDDING (2023): "Dynamic Spatial General Equilibrium," *Econometrica*, 91, 385–424.

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

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 (2024): "Information Technology and Returns to Scale," *American Economic Review*.

LIAO, W.-C. (2012): "Inshoring: The Geographic Fragmentation of Production and Inequality," *Journal of Urban Economics*, 72, 1–16.

MARTELLINI, P. (2022): "Local Labor Markets and Aggregate Productivity," Tech. rep., Working paper.

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

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

NAGY, D. K. (2023): "Hinterlands, City Formation and Growth: Evidence from the US Westward Expansion," *Review of Economic Studies*, 90, 3238–3281.

OBERFIELD, E. AND D. RAVAL (2021): "Micro Data and Macro Technology," *Econometrica*, 89, 703–732.

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

RAVIKUMAR, B., A. M. SANTACREU, AND M. SPOSI (2019): "Capital Accumulation and Dynamic Gains from Trade," *Journal of International Economics*, 119, 93–110.

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., K. GENADEK, R. GOEKEN, J. GROVER, AND M. SOBEK (2017): "Integrated Public Use Microdata Series: Version 7.0," *Minneapolis: University of Minnesota, 2017. https://doi.org/10.18128/D010.V7.0*.

SAIZ, A. (2010): "The Geographic Determinants of Housing Supply," *The Quarterly Journal of Economics*, 125, 1253–1296.

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.

SMITH, A. (1993): "Estimating Nonlinear Time-series Models using Simulated Vector Autoregressions," *Journal of Applied Econometrics*, 8, S63–S84.

——— (2008): "Indirect Inference," *The New Palgrave Dictionary of Economics, 2nd Edition*.

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*.

WALSH, C. (2023): "The Startup Multiplier," *Available at SSRN 3496782*.

Online Appendix Material for

Urban-Biased Growth: A Macroeconomic Analysis

by Fabian Eckert, Sharat Ganapati, and Conor Walsh

April 2024

FOR ONLINE PUBLICATION

A. Proofs and Derivations

In this section, we derive the main theoretical results of the paper.

A.1 Derivation of the Demand System

Let a intermediate input varieties be indexed by ω. Denote the total number of varieties in each sector by N_s. The representative firm producing the final good has the following production function:

$$Y = \left(\sum_s \left(\int_0^{N_s} q_s^{\zeta_s}(\omega) d\omega \right)^{\frac{\varsigma}{\zeta_s}} \right)^{\frac{1}{\varsigma}} := \left(\sum_s Q_s^{\varsigma} \right)^{\frac{1}{\varsigma}},$$

where the elasticity of substitution over firm varieties within a sector is $\iota_s := \frac{1}{1-\zeta_s}$ and the elasticity of substitution over sectoral CES bundles, Q_s, is $\gamma := \frac{1}{1-\varsigma}$. Solving the representative firm's profit maximization problem yields the standard demand curve for an individual variety:

$$p_s(\omega) = q_s(\omega)^{-\frac{1}{\iota_s}} \mathcal{D}_s \quad \text{where} \quad \mathcal{D}_s := \frac{P_s^{\frac{\iota_s - \gamma}{\iota_s}}}{P^{\frac{1-\varsigma}{\iota_s}}} I^{\frac{1}{\iota_s}},$$

and I denotes total demand for the final good. The optimal sectoral price index, P_s is defined by $P_s^{1-\iota_s} = \int_0^{N_s} p_s(\omega)^{1-\iota_s} d\omega$ and the ideal price index of the final good, P, is defined by $P^{1-\gamma} = \sum_s P_s^{1-\gamma}$. The term $\mathcal{D}_s$ is a measure of sector-specific aggregate demand. Using the preceding expression, we can then express a firm's revenue function in terms of the output or the price of its variety:

$$\text{(OA.1)} \qquad r_s(\omega) := y(\omega)^{\zeta_s} \mathcal{D}_s \quad \text{and} \quad r_s(\omega) := p(\omega)^{\frac{\zeta_s}{\zeta_s - 1}} \mathcal{D}_s^{\frac{1}{1-\zeta_s}}.$$

A.2 Proof of Theorem 1

We first state Lemma that links the rental rate and the investment price of capital in the steady state.

Lemma 1. *In steady state, the following holds for the rental rate of capital type f:*

$$w_{ft}^K = (\bar{R} - (1 - \delta_f^K)) p_{ft}^K = (\bar{R} - (1 - \delta_f^K)) \frac{1}{\mathcal{Z}_f},$$

where $\bar{R}$ comes from the steady-state Euler equation, and is invariant at $1/\beta$.

Proof. The rate of return of a unit of type-f capital in any period t is

$$R_t = \frac{w^K_{ft}}{p^K_{ft}} + \frac{p^K_{ft+1}}{p^K_{ft}}(1-\delta^K_f),$$

Note that in steady state the investment price of capital is constant over time, but then $p^K_{ft} = p^K_{ft+1} = 1/\mathcal{Z}_f$. Noting that in steady state $R_t = \bar{R} = 1/\beta$, then yields the result. □

Definition 1. *In steady state, let the present-discounted value of the expected lifetime payments of a firm in sector s to factor f be denoted by $\Phi_{\ell sf}$, so that:*

$$\Phi_{\ell sf} := \frac{\partial e_{\ell s}(\boldsymbol{w}_{\ell s})}{\partial w_{\ell sf}} w_{\ell sf} + \kappa \int \frac{\partial c_{\ell s}(z; y, \boldsymbol{w}_{\ell s})}{\partial w_{\ell sf}} w_{\ell sf} d\Omega_s(z),$$

where $\kappa := (\beta+1)/(\beta\xi+1) > 1$. Also let the total present-discounted value of the expected lifetime payments of a firm in sector s to all factors f of a particular type X be denoted by $\Phi^X_{\ell s}$, so that:

$$\Phi^X_{\ell s} := \sum_{f \in \mathcal{F}^X} \Phi_{\ell sf} \quad \textit{where} \quad X = K, L, M.$$

Definition 2. *Let the cost-share-weighted log change in the average price rate of all local factors be given by*

$$d\log \bar{w}_{\ell s} := \sum_{f \in \mathcal{F}^L \cup \mathcal{F}^M} \phi_{\ell sf} d\log w_{\ell sf} \quad \textit{where} \quad \phi_{\ell sf} := \frac{\Phi_{\ell sf}}{\sum_{f' \in \mathcal{F}^L \cup \mathcal{F}^M} \Phi_{\ell sf'}}.$$

For convenience, we restate Theorem 1 here and then prove it.

Theorem. *In the steady state, the general equilibrium response of average local-factor prices in a location-sector to a change in the investment price of type-f capital, p^K_f, is given by*

$$d\log \bar{w}_{\ell s} = -\frac{\Phi^K_{\ell sf}}{\Phi^L_{\ell s} + \Phi^M_{\ell s}} d\log p^K_f + \frac{\Phi^L_{\ell s} + \Phi^M_{\ell s} + \Phi^K_{\ell s}}{\Phi^L_{\ell s} + \Phi^M_{\ell s}} d\log \mathcal{D}_s.$$

Proof. Consider the free-entry condition in steady state:

$$\begin{aligned} e_{\ell s}(\mathbf{w}_{\ell s}) &= \kappa \int \pi_{\ell s}(z) d\Omega_s(z) \\ &= \kappa \int \max_y [y^{\zeta_s} \mathcal{D}_s - z^{-1} y v_{\ell s}(y, \mathbf{w}_{\ell s})] d\Omega_s(z). \end{aligned}$$

Now by the envelope theorem:

$$\frac{\partial \pi_{\ell s}(z)}{\partial w_{\ell s f}} = -z^{-1} y_\star \frac{\partial v_{\ell s}(y, \mathbf{w}_{\ell s})}{\partial w_{\ell s f}}.$$

where $y_\star$ denotes the profit-maximizing level of output of them firm. In addition, we also have

$$\frac{\partial \pi(z)}{\partial \mathcal{D}_s} = y_\star^{\zeta_s},$$

Totally differentiate the free-entry condition and use these expressions to obtain

(OA.2)

$$\sum_f \frac{\partial e_{\ell s}(\mathbf{w}_{\ell s})}{\partial w_{\ell s f}} dw_{\ell s f} = \kappa \int \left[y_\star^{\zeta_s} D_s d\log \mathcal{D}_s - \sum_f z^{-1} y_\star \frac{\partial v_{\ell s}(y_\star, \mathbf{w}_{\ell s})}{\partial w_{\ell s f}} dw_{\ell s f} \right] d\Omega_s(z).$$

We can also write the free-entry condition using Shephard's Lemma as

$$\text{(OA.3)} \qquad \sum_f \frac{\partial e_{\ell s}(\mathbf{w}_{\ell s})}{\partial w_{\ell s f}} w_{\ell s f} = \kappa \int \left[y_\star^{\zeta_s} D_s - \sum_f z^{-1} y_\star \frac{\partial v_{\ell s}(y_\star, \mathbf{w}_{\ell s})}{\partial w_{\ell s f}} w_{\ell s f} \right] d\Omega_s(z).$$

Using equations (OA.3) in (OA.2) yields

$$\sum_f \left[\frac{\partial e_{\ell s}(\mathbf{w}_{\ell s})}{\partial w_{\ell s f}} w_{\ell s f} + \kappa \int z^{-1} y_\star \frac{\partial v_{\ell s}(y_\star, \mathbf{w}_{\ell s})}{\partial w_{\ell s f}} w_{\ell s f} d\Omega_s(z) \right] d\log w_{\ell s f}$$
$$= \sum_f \left[\frac{\partial e_{\ell s}(\mathbf{w}_{\ell s})}{\partial w_{\ell s f}} w_{\ell s f} + \kappa \int z^{-1} y_\star \frac{\partial v_{\ell s}(y_\star, \mathbf{w}_{\ell s})}{\partial w_{\ell s f}} w_{\ell s f} d\Omega_s(z) \right] d\log \mathcal{D}_s.$$

We use Definition 1 to simplify the preceding equation:

$$\sum_f \Phi_{\ell s f} d\log w_{\ell s f} = \sum_f \Phi_{\ell s f} d\log \mathcal{D}_s.$$

Next, note that by Lemma 1, the steady state change in the rental rate of type-f capital is independent of the rental rate of type-f' capital. As a result, in response to the change of the investment price in type-f capital, we have $d\log w_{\ell s f} \neq 0$ and $d\log w_{\ell s f'} = 0 \forall f' \in \mathcal{F}^K \setminus f$.

Using this and Definition 2, we can write:

$$d\log \bar{w}_{\ell s} = -\frac{\Phi^K_{\ell s f}}{\Phi^L_{\ell s} + \Phi^M_{\ell s}} d\log w^K_f + \left(\frac{\Phi^L_{\ell s} + \Phi^M_{\ell s} + \Phi^K_{\ell s}}{\Phi^L_{\ell s} + \Phi^M_{\ell s}}\right) d\log \mathcal{D}_s.$$

Lastly, we can use Lemma 1 to replace the rental rate of capital with its investment price, so that:

$$d\log\bar{w}_{\ell s} = -\frac{\Phi^K_{\ell sf}}{\Phi^L_{\ell s}+\Phi^M_{\ell s}} d\log p^K_f + (\frac{\Phi^L_{\ell s}+\Phi^M_{\ell s}+\Phi^K_{\ell s}}{\Phi^L_{\ell s}+\Phi^M_{\ell s}}) d\log\mathcal{D}_s,$$

which concludes the proof. □

A.3 Derivation of Equations in the Simple Model

We derive the expressions in the simple model section in a number of steps.

Profit Maximization and Free Entry Recall that in the simple model, firms do not differ in productivity; without loss of generality, we set $z = 1$ for all firms. We can write a firm's variable profits as follows:

$$\pi_\ell = py - yA_\ell^{-1}v(y, \mathbf{w}_\ell)$$

Taking first-order condition with respect to output and re-arranging yields:

$$p_\ell = \frac{1}{\zeta}A_\ell^{-1}v(y, \mathbf{w}_\ell)(1 + \frac{\partial \log v(y, \mathbf{w}_\ell)}{\partial \log y}) := \frac{1}{\zeta}A_\ell^{-1}v(y, \mathbf{w}_\ell)(1 + \bar{v}_\ell),$$

where $\bar{v}_\ell$ is a measure of the increasing returns to scale induced by the non-homotheticity in production. Note that $\bar{v}_\ell = 0$ in the homothetic case.

We can plug the optimal pricing expression into the definition for aggregate demand to rewrite the expression for aggregate demand in equation (OA.1) in terms of factors prices and output. Combining the pricing rule with the demand function,

$$y^{\zeta-1}\mathcal{D} = \frac{1}{\zeta}A_\ell^{-1}v(y, \mathbf{w}_\ell)(1 + \bar{v}_\ell). \tag{OA.4}$$

Plugging the pricing rule into the expression of firm variable profits and setting the result equal to the entry cost yields

$$\pi^\star_\ell = yA_\ell^{-1}v(y, \mathbf{w}_\ell)(\frac{1-\zeta}{\zeta} + \frac{1}{\zeta}\bar{v}_\ell) = e(\mathbf{w}_\ell). \tag{OA.5}$$

Equations (OA.4) and (OA.5) pin down wages and output, given capital rental rates and aggregate demand, making wages and output are the only endogenous variables that vary in the cross-section of locations.

Exposure in the Cross-Section In the simple model, the exposure elasticity simplifies to the following:

$$\Lambda_\ell = \frac{w^K X_\ell^K}{w_\ell^L X_\ell^L} = \frac{w^K}{w_\ell^L} \frac{y A_\ell^{-1} v_{w^K} + e_{w^K}}{y A_\ell^{-1} v_{w_\ell^L} + e_{w^L}}, \tag{OA.6}$$

where v_x and e_x for $x = w_\ell^L, w^K$ denote the derivative of the variable cost and entry-cost function with respect to the respective price. Differentiating equation (OA.6) in the cross-section of locations yields

$$\begin{aligned} d\log\Lambda_\ell = {} & \frac{yA_\ell^{-1}v_{w^K}}{yA_\ell^{-1}v_{w^K}+e_{w^K}} d\log y - \frac{yA_\ell^{-1}v_{w^K}}{yA_\ell^{-1}v_{w^K}+e_{w^K}} d\log A_\ell \\ & + \frac{yA_\ell^{-1}v_{w^K w_\ell^L}+f_{pw}}{yA_\ell^{-1}v_{w^K}+e_{w^K}} w_\ell d\log w_\ell + \frac{yA_\ell^{-1}v_{py}}{yA_\ell^{-1}v_{w^K}+e_{w^K}} y d\log y \\ & - d\log w_\ell - \frac{yA_\ell^{-1}v_{w_\ell^L}}{yA_\ell^{-1}v_{w_\ell^L}+e_{w_\ell^L}} d\log y - \frac{yA_\ell^{-1}v_{w_\ell^L y}}{yA_\ell^{-1}v_{w_\ell^L}+e_{w_\ell^L}} d\log y \\ & + \frac{yA_\ell^{-1}v_{w_\ell^L}}{yA_\ell^{-1}v_{w_\ell^L}+e_{w_\ell^L}} d\log A_\ell - \frac{yA_\ell^{-1}v_{w_\ell^L w_\ell^L}+f_{w_\ell^L w_\ell^L}}{yA_\ell^{-1}v_{w_\ell^L}+e_{w_\ell^L}} w_\ell^L d\log w_\ell^L. \end{aligned}$$

We define the following cost shares:

$$\theta_\ell^{V|L} := \frac{yA_\ell^{-1}v_{w_\ell^L}}{yA_\ell^{-1}v_{w_\ell^L}+e_{w_\ell^L}} \quad \text{and} \quad \theta_\ell^{V|K} := \frac{yA_\ell^{-1}v_{w^K}}{yA_\ell^{-1}v_{w^K}+e_{w^K}},$$

which give the fraction of total labor (capital) payments that go to variable cost as opposed to entry costs. Using this definition, we rearrange terms to derive the expression from the body of the paper:

$$\frac{d\log\Lambda_\ell}{d\log A_\ell} = (\sigma_{KL}-1)\frac{d\log w_\ell^L}{d\log A_\ell} + \epsilon_{KL}\frac{d\log y}{d\log A_\ell} + (\theta^{V|K}-\theta^{V|L})\frac{d\log y/A_\ell}{d\log A_\ell} \tag{OA.7}$$

where

$$\sigma_{KL} := \frac{\partial\log K_\ell/L_\ell}{\partial\log w_\ell^L/w^K} \quad \text{and} \quad \epsilon_{KL} := \frac{\partial\log K_\ell/L_\ell}{\partial\log y},$$

are the elasticity of substitution between capital and labor for a given level of output and the capital intensity in variable cost as a function of firm scale, for given factor prices. K_ℓ and L_ℓ are the total amounts of capital and labor used in location ℓ across both variable and entry cost.

Next, we use equations (OA.4) and (OA.5) to find expression for the cross-sectional terms in equation (OA.7) in the equilibrium of the model. First, differentiate equation (OA.4) in the cross-section of locations to obtain

$$(\zeta - 1) d \log y = -d \log A_\ell + \frac{v_{w_\ell^L}(y, \mathbf{w}_\ell) w_\ell^L}{v(y, \mathbf{w}_\ell)} d \log w_\ell^L + \frac{v_y(y, \mathbf{w}_\ell)}{v(y, \mathbf{w}_\ell)} y d \log y. \tag{OA.8}$$

We define the following two cost shares:

$$\theta_\ell^{L|V} := \frac{v_{w_\ell^L}(y, \mathbf{w}_\ell) w_\ell^L}{v(y, \mathbf{w}_\ell)} \quad \text{and} \quad \theta_\ell^{L|E} := \frac{e_{w_\ell^L}(\mathbf{w}_\ell) w_\ell^L}{e(\mathbf{w}_\ell)},$$

which denote the share of labor payments in variable and entry costs, respectively. Using this definition, we simplify equation (OA.8) to:

$$[(\zeta - 1) - \bar{v}_\ell] d \log y = \theta_\ell^{L|V} d \log w_\ell^L - d \log A_\ell. \tag{OA.9}$$

Now, we differentiate equation (OA.5) in the cross-section of locations to obtain:

$$y A_\ell^{-1} v(y, \mathbf{w}_\ell) [\frac{1-\zeta}{\zeta} + \frac{1}{\zeta} \bar{v}_\ell] d \log y - y A_\ell^{-1} v(y, \mathbf{w}_\ell) [\frac{1-\zeta}{\zeta} + \frac{1}{\zeta} \bar{v}_\ell] d \log A_\ell$$
$$+ y A_\ell^{-1} v_{w_\ell^L}(y, \mathbf{w}_\ell) [\frac{1-\zeta}{\zeta} + \frac{1}{\zeta} \bar{v}_\ell] w_\ell^L d \log w_\ell^L = e_{w_\ell^L}(\mathbf{w}_\ell) w_\ell^L d \log w_\ell^L,$$

which simplifies to

$$d \log y - d \log A_\ell = [\theta_\ell^{L|E} - \theta_\ell^{L|V}] d \log w_\ell^L \tag{OA.10}$$

Plugging (OA.9) into equation (OA.10) and rearranging, we find:

$$\frac{d \log w_\ell^L}{d \log A_\ell} = \frac{(\zeta - 1) - \bar{v}_\ell + 1}{\theta_\ell^{L|V} - [(\zeta - 1) - \bar{v}_\ell][\theta_\ell^{L|E} - \theta_\ell^{L|V}]}. \tag{OA.11}$$

Plugging this back into equation (OA.9) yields:

$$\frac{d \log y}{d \log A_\ell} = \frac{\theta_\ell^{L|E}}{\theta_\ell^{L|V} - [(\zeta - 1) - \bar{v}_\ell][\theta_\ell^{L|E} - \theta_\ell^{L|V}]}. \tag{OA.12}$$

We now set $\bar{v}_\ell = 0$ and combine equations (OA.11) and (OA.12) to obtain the following two cross-sectional relationships:

$$\frac{d \log w_\ell}{d \log A_\ell} = \frac{\zeta}{\zeta \theta_\ell^{L|V} + (1 - \zeta) \theta_\ell^{L|E}} \quad \text{and} \quad \frac{d \log y}{d \log A_\ell} = \frac{\theta_\ell^{L|E}}{\zeta \theta_\ell^{L|V} + (1 - \zeta) \theta_\ell^{L|E}} \tag{OA.13}$$

which are the two expressions that appear in the expression for how the exposure elasticity varies with location productivity.

Figure 7 in the body of the paper also shows some special cases of the preceding expressions. We now discuss a set of special cases including the ones shown in the paper.

No Labor in Entry Cost. In this case, $\theta_\ell^{L|E} = 0$ and equation (OA.13) simplify to

$$\frac{d\log w_\ell}{d\log A_\ell} = 1/\theta_\ell^{L|V} \quad \text{and} \quad \frac{d\log y}{d\log A_\ell} = 0.$$

If labor is not used for entry costs, entry costs cannot vary across locations. As a result, firms have to be equally profitable in all locations in equilibrium. Therefore labor costs have to increase with local productivity to offset the entire productive advantage in equilibrium.

Only Labor in Entry Cost. In this case, $\theta_\ell^{L|E} = 1$ and equation (OA.13) simplify to

$$\frac{d\log w_\ell}{d\log A_\ell} = \frac{\zeta}{\zeta\theta_\ell^{L|V} + (1-\zeta)} \quad \text{and} \quad \frac{d\log y}{d\log A_\ell} = \frac{1}{\zeta\theta_\ell^{L|V} + (1-\zeta)}.$$

With only labor in the entry cost, entry costs increase one-for-one with wages. In more productive locations, firms are more productive and hence sell larger quantities. Wages increase less than output because the curvature of demand depresses the marginal product of labor at higher levels of output.

Cobb-Douglas Case: ζ=0 In this case, $\zeta \to 0$ and equation (OA.13) simplify to

$$\frac{d\log w_\ell}{d\log A_\ell} = 0 \quad \text{and} \quad \frac{d\log y}{d\log A_\ell} = 1.$$

If the demand system is Cobb-Douglas, the representative firm spends a fixed share of its expenditure on each firm. As a result, the revenue of each firm is invariant. Firms in more productive locations produce larger quantities but at lower prices, so that the marginal product of workers is the same in all locations regardless of their productivity.

Linear Production Function: ζ=1 In this case, $\zeta \to 1$ and equation (OA.13) simplify to

$$\frac{d\log w_\ell}{d\log A_\ell} = 1/\theta_\ell^{L|V} \quad \text{and} \quad \frac{d\log y}{d\log A_\ell} = \theta_\ell^{L|E}/\theta_\ell^{L|V}.$$

Entry costs are higher in more productive locations to offset higher profitability through higher productivity. In this case, the marginal product of labor does not fall as the firm increases output, so wages can increase in inverse proportion to their cost share with productivity. Output is higher in more productive locations. The more so the higher the labor share in entry costs because a larger labor share implies that entry costs increase more steeply with location productivity.

A.4 Factor Demands in the Quantitative Model

Consider the production technology from equation (11) that is common to all firms i. Solving the cost minimization problem of a firm then yields the following cost function:

$$c(z;y,\mathbf{w}_{\ell s}) = yz^{-1}\left[\left[(w^h_{\ell s})^{1-\sigma_s}A^h_{\ell s}y^{\bar{\epsilon}_s} + (w^k)^{1-\sigma_s}A^k_s\right]^{\frac{1-\varphi_s}{1-\sigma_s}} + (w^l_{\ell s})^{1-\varphi_s}A^l_{\ell s}\right]^{\frac{1}{1-\varphi_s}}.$$

Using Shepard's lemma, we derive the following expressions for the individual factor demands:

$$\begin{aligned} h &= c(z;y,\mathbf{w}_{\ell s})^{\varphi_s}P_x^{\sigma_s-\varphi_s}(w^h_{\ell s})^{-\sigma_s}A^h_{\ell s}y^{\bar{\epsilon}_s} \\ k &= c(z;y,\mathbf{w}_{\ell s})^{\varphi_s}P_x^{\sigma_s-\varphi_s}(w^k)^{-\sigma_s}A^k_s \\ l &= c(z;y,\mathbf{w}_{\ell s})^{\varphi_s}(w^l_{\ell s})^{-\varphi_s}A^l_{\ell s}, \end{aligned}$$

where $P_x^{1-\sigma_s} = (w^h_{\ell s})^{1-\sigma_s}A^h_{\ell s}y^{\bar{\epsilon}_s} + (w^k)^{1-\sigma_s}A^k_s$. Using these factor demands, we can find the following input ratios:

$$\frac{h}{k} = \frac{(w^h_{\ell s})^{-\sigma_s}A^h_{\ell s}y^{\bar{\epsilon}_s}}{(w^k)^{-\sigma_s}A^k_s} \quad \text{and} \quad \frac{h}{l} = \frac{P_X^{\sigma_s-\varphi_s}(w^h_{\ell s})^{-\sigma_s}A^h_{\ell s}y^{\bar{\epsilon}_s}}{(w^l_{\ell s})^{-\varphi_s}A^l_{\ell s}} \quad \text{and} \quad \frac{k}{l} = \frac{P_X^{\sigma_s-\varphi_s}(w^k)^{-\sigma_s}A^k_s}{(w^l_{\ell s})^{-\varphi_s}A^l_{\ell s}}.$$

With these expression, computing closed form expressions for scale-elasticities is straightforward, we find:

$$\frac{\partial \log \frac{k}{h}}{\partial \log y} = -\bar{\epsilon}_s;\ \frac{\partial \log \frac{h}{l}}{\partial \log y} = \bar{\epsilon}_s[1 - \frac{\varphi_s-\sigma_s}{1-\sigma_s}]\theta_{\ell s}(w_{\ell s},y);\ \frac{\partial \log \frac{k}{l}}{\partial \log y} = -\bar{\epsilon}_s\frac{\varphi_s-\sigma_s}{1-\sigma_s}\theta_{\ell s}(w_{\ell s},y),$$

where $\theta_{\ell s}(w_{\ell s},y) \in (0,1)$ is the cost share of skilled labor in the capital-skill bundle.

A.5 Endogenous Local Fundamentals

A long literature suggests 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 $B_{\ell h} = \bar{B}_{\ell h}\phi_\ell^{\chi_1}$, where ϕ_ℓ is the ratio of college- to non-college-educated workers in location ℓ. We borrow the parameter $\chi_1 = 2.6$ 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 ($\bar{B}_{\ell h}$). Column 5 of Table 4 presents the resulting wage-density gradients in 2015.

Ahlfeldt and Pietrostefani (2019) provide estimates for productivity spillovers from a meta-study of urban economics papers. We change the specification of local labor productivity shifters in our model for workers of type f as follows:

$$A^f_{\ell s} = \bar{A}^f_{\ell s}(X^L_\ell)^{\chi_2},$$

where X^L_ℓ indicates the total population count in location ℓ. The study by Ahlfeldt and Pietrostefani (2019) implies $\chi_2 = 0.04$. Column 6 of Table 4 presents the resulting wage-density gradients in 2015.

B. Data Construction

In this section, we provide additional details on the datasets used in the body of the paper.

B.1 Longitudinal Business Database (LBD)

We use the administrative, establishment-level LBD data from the US Census Bureau from 1980-2015. The LBD reports industry codes for establishments 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 received further updates in subsequent years. We use Fort and Klimek (2016) to crosswalk historical SIC information into consistent NAICS records. We trim outlier data, remove establishments without employment or payroll data, and omit establishments with mean worker pay greater than $1,000,000 per year.

The LBD also contains information on which firm owns each establishment, allowing us to combine it with other US Census datasets that report information on US firms.

B.2 Annual Capital Expenditures Survey (ACES)

The ACES provides broad-based statistics on business spending for new and used structures and equipment. United States Code, Title 13, authorizes this survey and provides for mandatory responses. Supplemental to the current Annual Capital Expenditure Survey, the Information and Communication Technology Survey (ICTS) collects data on non-capitalized and capitalized business spending for information and communication technology (ICT) equipment.

The ICTS covers all domestic, private, and non-farm firms. The ICTS sample consists of approximately 46,000 companies with one or more employees. Larger companies are selected yearly from the updated Business Register (BR); the survey includes all companies with at least 500 paid employees. Smaller companies with employees are stratified by industry and payroll size and selected randomly by strata.

The survey includes four types of ICT equipment and software: computer and peripheral equipment; ICT equipment excluding computers and peripherals; electromedical and electrotherapeutic apparatus; and computer software. Companies report non-capitalized and capitalized expenses.

Data reporting changed with the 2013 survey, the first for which firms reported electronically. The Census used mail-out/mail-back survey forms to collect data in previous survey years. As a result, our analysis relies mainly on the 2013 iteration of the survey. After 2013, the Census ran out of funding for the ICTS and discontinued it.

For 2013, we merged our LBD data with the ICTS data using the firm identifiers provided in both surveys. We excluded electromedical and electrotherapeutic apparatus from our analysis and aggregated all IT assets into a single measure of IT capital for each firm in the survey. Using this information, we constructed the measure of IT expenditure per worker described in the paper's body. Our results are robust to using earlier survey years.

B.3 US Decennial Census and American Community Survey

The LBD data does not contain information on the workers at each establishment. We create an additional panel dataset using information from the 1970, 1980, 1990, and 2000 US Decennial Census and the 2010 and 2015 American Community Survey (Ruggles et al., 2017). The panel contains total employment and labor income for each commuting zone, NAICS 1 sector, education group, occupation group, and year.

In constructing the panel from microdata, 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 (Tolbert and Sizer, 1996) covering the entirety of the continental US. 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. For 1970 and 1980, the crosswalk uses Census "county groups" instead of PUMA identifiers.

We aggregate all our data into 1-digit NAICS sectors designed to capture the principal functional differences between industry groups. 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 log 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. Figure OA.15 shows the resulting rent-price index for 1980 and 2015.

B.4 Quarterly Census of Employment and Wages (QCEW)

For some of our aggregate wage, employment, and establishment statistics (such as Figures OA.4 and 8), we use the publicly-available QCEW published by the Bureau of Labor Studies. The data come from unemployment insurance records and cover most US workers. 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 used the SIC industry classification standard. To convert this to the modern NAICS industry standard, we 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 classified ″SIC 1520″ as a non-Business Services industry and ″SIC 9999″ (non-classifiable establishments) as a non-Business Services industry.

B.5 Current Population Survey (CPS)

We obtain information on employee characteristics by firm size from the CPS conducted by the US Census Bureau. We accessed the data via IPUMS (Ruggles et al., 2017). Since 1992, the CPS has consistently asked respondents to report the size of their employer using the following bins: ″<10 employees″, ″10-24″, ″25-99″, ″100-499″, ″500-999″, and ″1000+.″ Data on employer size started in 1988; however, employer-size bins changed several times in the first few years of coverage. The question reached its current form in 1992, so we use that year in our calibration. We drop employees working more than 168 hours per week and part-time workers who worked less than 30 hours in a ″usual″ week. We classify workers with more than a bachelor′s degree as ″college-educated″ and all other workers as ″non-college.″

B.6 County Business Patterns (CBP)

As a robustness exercise, we document the increase in the wage-density gradient in the US Census Bureau's CBP database in Figure OA.5. The CBP provides total payroll and employment for each US county from 1980-2015.

We perform minimal processing of the data. We aggregate counties to commuting zones following Tolbert and Sizer (1996). We compute total payroll and total employment for each commuting zone and compute average wages as their ratio. We deflate average wages using the BEA PCE Deflator.

B.7 Bureau of Economic Activity (BEA) Fixed-Asset, Investment, and Value-Added Data

We use the BEA Fixed Asset Tables' "Detailed Data for Fixed Assets and Consumer Durable Goods" as our source of aggregate information on capital stocks and capital investments by sector.

Our first and most direct output from the BEA data is a set of capital-type-specific price indices. In particular, we extract the price indices for equipment capital and its subcategories: information processing, industrial equipment, transportation equipment, and other equipment. Similarly, we extract the price indices for intellectual property capital and its subcategories: software, research and development, and entertainment.[50] The left panel of Figure OA.13 shows the equipment capital price series, and the right panel for intellectual property. The most important takeaway from these figures is that most of the decline in equipment and intellectual property capital investment price is due to information processing equipment and software.

Next, we extract several data series for more granular asset categories. In particular, we extract the following information: (1) capital stock data in dollars, (2) capital quantity index, (3) capital investment information, and (4) capital depreciation rates.[51] We obtain these information for the following assets that we jointly define as "IT assets:" ENS1: Prepackaged software; ENS2: Custom software; ENS3: Own account software; EP1A: Mainframes; EP1B: PCs; EP1C: DASDs; EP1D: Printers; EP1E: Terminals; EP1F: Tape drives; EP1G: Storage devices; EP1H: System integrators; EP12: Office and accounting equipment; EP31: Photocopy and related equipment.

We compute a capital price series for each asset by dividing its nominal stock by the

[50] The nine series we extract have the following numbers in the BEA tables: Y033RG3Q086SBEA, Y034RG3Q086SBEA, A680RG3Q086SBEA, A681RG3Q086SBEA, A862RG3Q086SBEA, Y001RG3Q086SBEA, B985RG3Q086SBEA, Y006RG3Q086SBEA, Y020RG3Q086SBEA.

[51] We use the following data series respectively: (1) Current-Cost Net Capital Stock of Private Nonresidential Fixed Assets; (2) Fixed-Cost Net Capital Stock of Private Nonresidential Fixed Assets; (3) Investment in Private Nonresidential Fixed Assets; (4) Current-Cost Depreciation of Private Nonresidential Fixed Assets.

corresponding quantity index. Using this price index, we adjust the data on capital stocks and investments for each sector and year to be in 2015 dollars.

Our first output from the more granular data is the numbers on investment per worker across 1-digit NAICS sectors in 1980 and 2015. To construct this figure, we first aggregate investment in the more granular capital categories into investment in three broad types of IT capital: (1) proprietary software by combining counts for codes ENS2 and ENS3, (2) pre-packaged software simply as code ENS1, and (3) hardware by combining codes EP1A to EP31. We obtain employment for each sector and year from the QCEW. Figure 5 shows investment per worker across NAICS-1 sectors in 2015 dollars for our three categories of IT capital.

An important input into our calibration is the aggregate series of IT capital investment prices, which we use to calibrate the productivity of capital production in our model. Sectors differ in how much of each of the more granular IT capital assets they use at any point in time, whereas in our model, there is just one type of IT capital. To address this, we construct a sector-specific IT investment price. We take the price of the most granular assets in each class from the BEA. We then compute a sectoral ideal (Fischer) price index following the methodology of the BEA. The so-constructed price indices for each sector account for the difference in the composition of the IT capital bundle across sectors. Lastly, we deflate the sector-specific indices using the BEA PCE deflator. Figure OA.16 shows the resulting IT price index for both sectors; they are very similar.

We compute the average depreciation rate of the IT capital in each sector by weighting the depreciation rate for each asset type within the IT category by its stock in the sector. The result is a time-varying series of sector-specific depreciation rates for IT capital, which we feed directly into the model as its depreciation rate parameter. Like average IT prices, the depreciation rate of IT assets across sectors looks very similar.

To calibrate the productivity of IT capital in each sector, A_s^k, we target a measure of how much each sector spends on capital relative to labor. In particular, we sum nominal capital stocks each year into a single IT capital stock for Business Services and the rest of the economy. Then, we divided these stocks by the total payroll for each sector and year. Table OA.1 presents the results. The Business Services sector's IT capital stock per every dollar of payroll is significantly above that of the rest of the economy.

Note that the ratio in Table OA.1 corresponds to an aggregate version of the model-implied exposure elasticity for the case where capital depreciates fully each period and firms exit after one period.

TABLE OA.1: IT CAPITAL STOCK TO PAYROLL RATIO BY SECTOR AND YEAR

	1980	2015
Business Services	0.16	0.24
Other Sectors	0.05	0.07

Notes: We take IT Capital stocks data by sector in 1980 from the BEA Fixed Asset by Sector tables (see full details in Section B.7). We then use the QCEW to recover total payroll by sector in 1980.

C. ADDITIONAL FIGURES AND TABLES

In this section, we present additional figures and tables referenced in the main part of the paper. We provide more information on the data used in this section in Section B of the Online Appendix.

C.1 Supporting Evidence for Section 1.2

This section presents supporting evidence for the urban-biased growth phenomenon we documented in Section 1.2.

Urban-biased Growth in Other Datasets. Figure 2 in the paper's body used LBD data. Figure OA.1 replicates Figure 2 in other datasets. Panel A presents the result from the main paper using the LBD for ease of comparison. Panel B presents the result using data from the US Decennial Census. Using the Census somewhat attenuates our finding perhaps due to noise of self-reporting. Panel C presents the result in the QCEW, which is very similar to the LBD, but uses unemployment instead of tax records reported for on the worker instead of the establishment level; both datasets are of administrative quality. Lastly, Panel D presents the results in the County Business Patterns which are a public, tabulated version of the LBD data.

Spatially-biased Growth with other Commuting Zone Orderings. Figure 2 in the paper's body relied on ordering commuting zones by their 1980 population density and then grouping them into deciles of employment. Figure OA.2 replicates Figure 2 with alternative ways of constructing commuting zone deciles. Panel A shows the result when ordering by population density in the US Decennial Census. Panel B shows the result when ordering by total population size. Panel C shows the result when computing the density of a commuting zone as the tract-weighted population density. In constructing this alternative density measure, 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 (for example, the edges of the Los Angeles commuting zone). Finally, Panel D shows wage growth when ordering commuting zones by their average wage in 1980. Consistent with findings in Giannone (2022), wage growth appears flat when ordered by the initial wage of the commuting zone.

Urban-biased Growth before 1980. Figure 1 in the paper's body studied wage growth between 1980 and 2015 and showed that it was strongly urban-biased. We recreate Figure 1 with US Decennial Census/ACS data going back to 1950, since the LBD data is not available before 1975. Figure OA.3 shows that there was mildly urban-biased wage growth between 1950 and 1980, particularly in Business Services. However, the strongly urban-biased growth starting in 1980 presents a clear structural break. Our theory is consistent with urban-biased growth occurring before 1980. Because we are using the Decennial Census, the results for 1980 and 2015 do not exactly match those in Figure 1.

Using the Wage-Density Gradient to Measure Urban-biased Growth. In the paper's body, we show wage growth for deciles of commuting zones with increasing population density. An alternative way to document the urban bias in recent US wage growth is to study the changes in the relationship between average wages and population density over time. The so-called wage-density gradient describes the elasticity of wages to population density in the cross-section of US commuting zones. The wage-density elasticity is the preferred measure of the urban-wage premium among urban economists.

Figure OA.4 shows the evolution of the wage-density elasticity using data from the QCEW. This elasticity more than doubled between 1980 and 2008 before holding steady in the subsequent years, reflecting the urban-biased growth documented in Figure 2. In addition, we plot coefficients from quantile regressions of the same distribution and show that the wage-density gradient evolved similarly in all quantiles. Note that we recompute commuting zone density for each year in Figure OA.4.

The Evolution of the Wage-Density Gradient Across Datasets. Next, we demonstrate that the wage-density elasticity increases in all major data sets on the US labor market. Figure OA.5a shows the wage-density elasticity for each year computed in the QCEW, the LBD, the US Decennial Census, and the CBP. We provide information on these data sources in the data section of the Online Appendix. The wage-density coefficients in data from the QCEW, CBP, and LBD all have a similar level and show comparable trends over time. 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.

The Evolution of the Wage-Density Gradient for Different Density Measures. Next, we show that the wage-density elasticity has increased regardless of how we measure location density. In addition, we show similar results for the wage-population-size gradient. OA.5b shows the wage-density coefficient in the QCEW using different measures of commuting zone density. First, we show the gradient using the 1980 population density of a commuting zone for all years instead of recomputing density each year (cf. Figure OA.4). Second, we compute a commuting zone's employment density instead of its commuting zone density. Third, we use the 1980 tract-weighted density of a commuting zone. Finally, we show the wage-population elasticity instead of the wage-density elasticity, using 1980 commuting zone populations. All coefficients exhibit broadly similar trends.

The Evolution of the Wage-Density Across US Counties. Figure OA.5c shows 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 similar to that of the commuting zone estimates over time.

The Evolution of the Wage-Density Gradient in Europe. Figure OA.5d shows the wage-employment elasticity computed across locations within the EU-15 countries. Instead of wages, the outcome variable is GDP per worker. The regressor is employment instead of population density since we lack the area data for European locations. Europe shows trends similar to the US; the GDP-location size elasticity roughly doubles from about .04 in 1980 to about .08 in 2010.

The Evolution of the Wage-Density Gradient in the Microdata. Figure OA.6 shows the raw commuting-zone level data used to compute wage gradients in 1980 and 2015, within and outside the Business Services Sectors, in the Decennial Census and ACS.

C.2 Supporting Evidence for Section 1.3

This section presents additional figures and exhibits for Section 1.3 in the paper which introduces three facts on the urban-biased growth of the US economy.

Disaggregated Industry Detail within Sectors. The main decomposition in the paper in Figure 3 presents results for 1-digit NAICS sectors. Figure OA.7 replicates Figure 3 for 2-digit NAICS industries. The industries within Business Services that contribute most to urban bias are in descending order: professional services, finance, information, admin and waste, management, and real estate.

Employment and Wages at Large and Small Establishments. Figure 4 in the paper showed that wage growth at large Business Services establishments accounts for most urban-biased growth. In this section, we provide additional details. For disclosure reasons, for this section, we define Business Services as only 2-digit NAICS codes 51, 52, 54, 55, so we omit Real Estate and Administrative Services relative to the definition in the paper. Moreover, we define the large establishments as the largest establishments that jointly account for 50% of the US workforce in 2015 which leads to a cutoff of 200 workers; in the body of the paper we defined them in 1980.

With these caveats, Figure OA.9 shows employment and wages at large and small establishments across commuting zones, sectors, and decades. Panel A shows wages at large and small Business Services establishments in 1980 and 2015. Panel B shows wages at large and small establishments in all other sectors in 1980 and 2015. Panel C shows employment shares within each commuting zone decile at large and small Business Services establishments in 1980 and 2015. Panel D shows employment shares within each commuting zone decile at large and small establishments in other sectors in 1980 and 2015.

Figure OA.9 helps understand why wage growth at large Business Services establishments accounts for most urban-biased growth. Panel A shows large wage growth differences at large Business Service firms across commuting zones. At the same time Panel C shows that large Business Services establishments account for a larger employment share in high-density commuting zones. However, differences in employment shares across commuting zones are small compared to the wage growth differences in Panel A. Moreover, Panel B also shows that the cross-sectional patterns of employment shares are constant over time, so that differential changes in employment shares at large Business Services firms do not contribute to urban-biased growth.

Firms or Establishments and Urban-biased Growth. Figure 4 in the body of the paper shows the contribution of large and small establishments to the urban-biased growth of the US economy. To construct it, we first compute the size of the establishment that employed the median US worker in 1980 (about 100 workers) and then group establishments into those above and below this median. In this section, we instead compute the size of the firm that employed the median US worker in 1980 (about 1000 workers) and then group establishments into large and small based on whether the firm that controls them is above or below this median.

Using these two ways of defining large and small establishments, we present a similar decomposition as in Fact 2 in the body of the paper. In particular, we decompose the wage change in each location ℓ as follows:

$$\Delta w_\ell = \underbrace{\mu^L_{\ell O}\Delta w^L_{\ell O}}_{\text{OL}} + \underbrace{\mu^S_{\ell O}\Delta w^S_{\ell O}}_{\text{OS}} + \underbrace{\mu^L_{\ell N5}\Delta w^L_{\ell N5}}_{\text{N5L}} + \underbrace{\mu^S_{\ell N5}\Delta w^S_{\ell N5}}_{\text{N5S}} \quad + \underbrace{\sum_{se} w^e_{\ell s}\Delta\mu^e_{\ell s}}_{\text{S}} + \underbrace{\sum_{se} \Delta\mu^e_{\ell s}\Delta w^e_{\ell s}}_{\text{C}},$$

where $s = N5$ and $s = O$ denote the Business Services sector and other sectors, and $e = L$ and $e = S$ index large and small establishments, or establishments of large and small firms. $\mu^e_{\ell s}$ indicates the share of employment in location ℓ accounted for by type e establishments/firms in sector s. $w\ell s^e$ indicates the average of workers at type e establishments/firms in sector s in location ℓ. OL and OS refer to wage growth at large and small establishments 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 OA.10a presents the results of this decomposition and shows that most urban-biased wage growth occurred at establishments of large Business Services firms. Figure OA.10b replicates Figure OA.10a using the establishment-based size definition instead. The two figures are very similar, suggesting that establishments of large Business Services firms are themselves large establishments. Using establishments or firms as the unit of analysis does not affect the conclusion of our second fact: large establishments of large firms within Business Services drove the urban-biased growth of the US economy.

Aggregate Wage and Employment Growth in the Business Services Sector. The left panel of Figure OA.12 shows employment relative to 1980 for all NAICS-1 sectors in the US economy. We highlight employment in the Business Services sector in red. Business Services employment has more than doubled over this period. The only sector for which employment has decreased is manufacturing. Total US employment has approximately doubled in this period.

The right panel of Figure OA.12 shows average wages relative to 1980 for all NAICS-1 sectors in the US economy. We highlight wage growth in the Business Services sector in red. Business Services wages have almost doubled since 1980. In most other sectors, wage growth was below 40% over this period.

Overall, Figure OA.12 shows the explosive growth of the Business Services sector over our study period.

Information Technology Investments per Worker across 2-Digit NAICS Industries. Figure 5 in the paper showed information technology investments per worker for each 1-digit NAICS sector. Figure OA.8 replicates Figure 5 for 2-digit NAICS industries. Almost all sub-industries within the Business Services sector have made larger IT investments per worker than users of IT than any other industry in the US economy. Other industries that have made significant IT investments per worker are Natural Resources and Utilities in 1980 and Natural Resources, Utilities, and Wholesale in 2015 (see also Ganapati 2024).

Non-IT capital Investments per Worker across 1-Digit NAICS Industries. Figure 5 in the paper showed information technology investments per worker for each 1-digit NAICS sector. Figure OA.11 replicates Figure 5 but for investments in non-IT capital. Non-IT capital includes all private non-residential assets not classified as IT assets. In contrast with IT investments, the Business Services sector does not emerge as an outlier.

IT Expenditure per Worker in the Spiceworks Data. The main body of the paper uses data on firm-level IT expenditures from the Census ACES dataset. To our knowledge, the ACES is the only source of IT investments on the firm or establishment level, as provided by the US Census. To corroborate our evidence on IT expenditures across firms, we acquired a commercial data provider's additional dataset on IT investments across US establishments. The Spiceworks data was formerly known as Ci Technology

Database, produced by the Aberdeen Group, and before that as Harte-Hanks data. Due to its broad coverage and high accuracy, many prior academic publications in economics have used this data (e.g., Bresnahan et al., 2002; Beaudry et al., 2010; Bloom et al., 2016).

The Spiceworks data contains spending on different types of IT technologies for a large set of US establishments for several years. Spiceworks distributes a questionnaire to firms about their IT usage across their establishments, including their NAICS industry code, employment, location, and IT spending per location. We do our best to reconstruct the capital categories in the ACES in the Spiceworks data and use the 2015 data to be as close as possible to the year for which we used the ACES data, 2013. The advantage of the Spiceworks data is its broader coverage and the fact that it is collected at the establishment level, not at the firm level. As such, the Spiceworks data aligns more closely with the establishment focus of our analysis.

Table OA.4 replicates Table 1 using the Spiceworks data on the establishment level. We directly use an indicator for Business Services establishments and the establishment's location density instead of considering the share of each firm that works in Business Services or the average population density of a firm's establishments.[52] Employment reflects the total employment of a firm across establishments to align with our Census ACES data. Results are broadly consistent. However, coefficient magnitudes are slightly attenuated compared to the Census ACES data, perhaps reflecting measurement error or an unobserved imputation procedure.

Price Declines. Capital investment prices for equipment and intellectual property have declined dramatically since 1980. The left and right panel of Figure OA.13 shows the decline in the BEA price index for equipment and intellectual property between 1980 and 2018 in black, normalized by their 1980 levels, respectively. All price indices are relative to the BEA PCE deflator. Figure OA.13 also shows that the vast majority of the decline in both indices is due to declines in the price indices of information processing equipment (among equipment capital) and software (among intellectual property). Figure OA.13 shows why our paper focuses on IT capital, which combines information processing equipment and software: the investment prices of non-IT capital have moved very little since 1980, while the joint IT price index has declined dramatically.

C.3 Supporting Evidence for Section 1.4

The college share of employment in big cities increased sharply during the period under study (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-

[52]We had to use these proxy measures in the ACES data, which is only available at the firm level.

paying jobs. In particular, Business Services were already among the most skill-intensive sectors in the US economy in 1980, and they became even more skill-intensive by 2015. In this section, we explore the role of such compositional changes.

We use the Census data because the LBD data lacks demographic information. Because 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.[53] The last column of Table OA.2 presents the share of urban-biased growth accounted for by each 1-digit NAICS sector in the Census. Patterns are similar as in the corresponding decomposition in the LBD; see Figure 3: Business Services are by far the positive most significant contributor to urban-biased growth while the manufacturing sector is a large negative contributor; other sectors play virtually no role. Compared to the LBD data, the positive contribution of the Business Services sector in the Census data is more positive, and the contribution of trade and transport is negative but moderately so.

We introduce a decomposition of the difference in wage growth rates for 1980-2015 between high- and low-density commuting zones. The decomposition isolates a component of wage growth in each location and sector due to skill deepening alone. This ”skill-deepening” component captures the wage growth that would have resulted had the sector’s employment share and college wage premium remained constant at their 1980 values, but its college share of employment evolved as in the data.

To formalize this, we decompose wage growth in each location-sector similar to what we did in equation (1):

$$\delta_{\ell s} = \underbrace{\frac{\mu_{\ell st}(w^h_{\ell st} - w^l_{\ell st})\Delta\mu^h_{\ell st}}{\bar{w}_{\ell t}}}_{\text{Deepening}} + \zeta_{\ell s}, \tag{OA.14}$$

where $\Delta\mu^h_{\ell s}$ denotes the change in the share of employment in sector s in location ℓ accounted for by college-educated workers between t and $t+1$, $w^h_{\ell s}$ denotes average wages of college-educated workers in location ℓ and sector s, and $w^l_{\ell s}$ average wages of worker without a college degree. As before, we use equation OA.14 to compute the share of urban-biased growth due to differences in skill deepening in each sector across space and share due to differential changes in the residual component across regions.

The left two columns of Table OA.2 present the results from the decomposition in equation (OA.14) for college- versus non-college-educated workers. The changing

[53]In particular, workers in high-skill service firms that own manufacturing or retail establishments often misreport their sector as manufacturing or retail. For 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.

composition of urban economies toward more educated workers explains about 35% of urban-biased growth. Across sectors, the importance of education-deepening varies. Skill-deepening within Business Services explains only about 16.3% of the aggregate economy's urban-biased growth. At the same time, skill-deepening only explains slightly more than a tenth of all urban-biased growth in the Business Services sector.

Another recent line of work has studied the role of so-called cognitive non-routine (CNR) occupations in trends in aggregate and local inequality (see Rossi-Hansberg et al., 2019). We apply the same decomposition in equation (OA.14) to CNR and non-CNR workers instead of college and non-college workers. Columns 3 and 4 of Table OA.2 show occupational shifts within Business Services explain about 16.3% of all urban-biased growth, while the residual term, which captures within-occupation wage growth, explains the vast majority of urban-biased growth.

C.4 Additional Figures for Section 3

This section presents figures that show data moments that we use in the estimation of our model in Section 3.

High- to Low-Skill Ratio and Firm Size. The CPS routinely asks workers about the size of the firm they work for. We use this information to compute the college share across the firm-size distributions. Figure OA.14 shows the share of college-educated workers within the firm-size bins provided in the CPS data. In Business Services and other sectors, the college share of employment is higher at larger firms. However, the college share of employment for Business Services is about 15 percentage points higher for all firm sizes than the average college share in the other sectors. Separately for each sector, we use the coefficient on firm size in a regression of the log of the college share of employment on log firm size to calibrate φ_s, the elasticity of substitution between high and low-skill labor, as detailed in the text.

Commuting Zone Residential Rent Price Index. We construct a commuting zone rent index. We constructed the index using microdata on reported gross rents and dwelling characteristics from the US Census and ACS. We regressed the log of gross rents paid by individuals based on the building's age, the number of rooms, and a commuting-zone-year fixed effect. We interpret the commuting-zone-year fixed effect as a rent price index because it represents the price of a unit of observationally-equivalent housing in each commuting zone. The top two panels of Figure OA.15 show the rent index across commuting zones for 1980 and 2015.

C.5 Additional Figures for Section 4

This section presents exhibits that provide additional model outputs referenced as part of our urban growth accounting exercise in Section 4.

Residential Rent Prices Across Commuting Zones in the IT-Only Economy. The bottom left panel of Figure OA.15 shows residential rents across commuting zones in our calibrated model in 1980. The residential rents are exactly as in the data in the top left panel because we chose residential land supply to match the data on the commuting zone-level residential rent price index exactly.

The bottom right panel of Figure OA.15 shows residential rents across commuting zones in our counterfactual IT-only economy in 2015. The rent-density gradient increased markedly between the two years due to the decline in IT prices. In fact, in the IT-only economy, the 2015 rent-density gradient is steeper than in the data, suggesting that other forces unrelated to the IT capital investment price decline offset some of the gradient's steepening.

Large Firms and Urban-Biased Growth in the IT-Only Economy. Figure OA.17 replicates Figure 4 in its top panel and displays the corresponding decomposition in our counterfactual IT-only economy in the bottom panel. Large firms account for most urban-biased growth in our IT-only counterfactual, as in the data.

However, the model is less successful at replicating the split into wage growth versus employment growth in accounting for urban-biased growth. Relative to the data, a large part of the urban-biased growth at large Business Services firms in the counterfactual IT-only economy reflects differential employment growth across commuting zones rather than differential wage growth. In the model, all firms within a location-sector pay the same wage, while larger firms pay systematically higher wages in the data. Introducing firm-specific labor-supply curves would allow our model to more accurately capture the decomposition of large firm payroll growth into wage versus employment growth. However, introducing and estimating firm-specific labor supply curves is outside the scope of our paper.

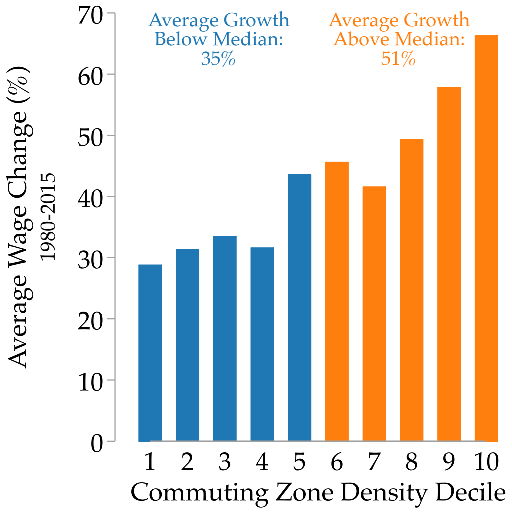
Calibrated Amenity Residuals across Commuting Zones and Education Groups. Figure OA.18 presents the calibrated location amenities in the model for 1980 and 2015, separately for college and non-college workers. For each education group, we normalized amenities by the value of amenities of the New York commuting zone in 1980.

Additional Accounting Results in the IT-Only Economy. In the body of the paper, we study how much of the increase in the wage-density gradient the IT-only economy exhibits. Table OA.3 presents additional results and robustness checks. Column "Only A" shows the resulting wage-density gradient if only fundamental productivities change from their calibrated values in 1980 to their calibrated values in 2015 while the IT price and all other structural residuals remain constant at their 1980 levels. The wage-density coefficient flattens due to rural-biased productivity growth for non-college-educated workers, seen in Figure 12. Column "Only B" shows the change in the wage-density elasticity if only amenities had changed from their calibrated values in 1980 to their calibrated values in 2015, while the IT price and all other structural residuals are held constant at their 1980 levels. It shows that changes in amenities had no impact on urban-biased growth.

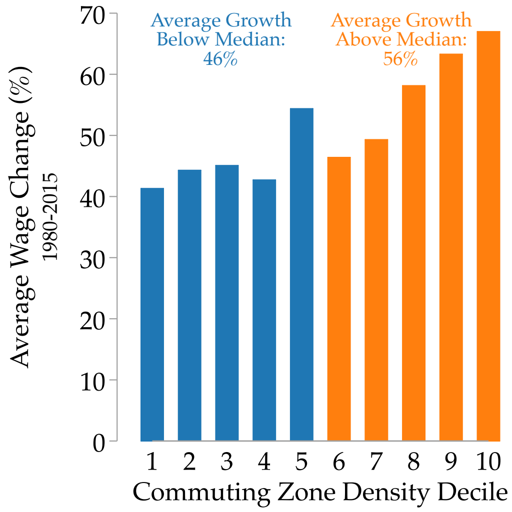
The remaining columns show how much urban-biased growth the IT-only economy can account for under alternative parameters. The Column "Low Elast." targets the macro elasticity of substitution between capital and labor from Oberfield and Raval (2021), which leads to a firm-level elasticity of substitution between skilled labor and capital of $\sigma_s = 0.2$. The Column "High Elast." targets the macro elasticity of substitution between capital and labor from Karabarbounis and Neiman (2014), which leads to a firm-level elasticity of substitution between skilled labor and capital of $\sigma_s = 0.6$. Finally, the Column "Equal Lab. Elast." sets labor-supply elasticities equal across education groups, with the spatial elasticity set to 4 and the sectoral elasticity set to 0.5, roughly averages of our estimated values for both elasticities.

FIGURE OA.1: WAGE GROWTH BY DENSITY: OTHER DATASETS

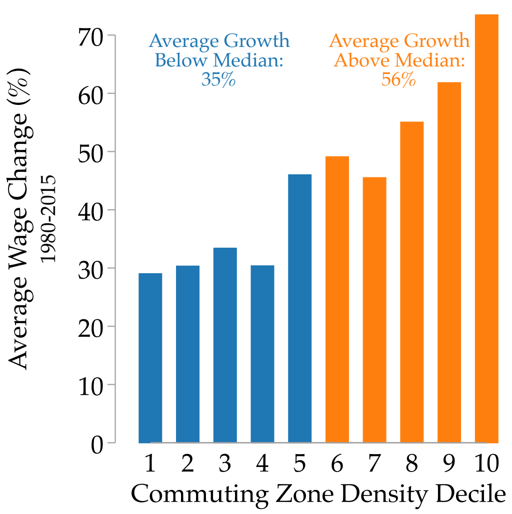
(A) LBD-Baseline



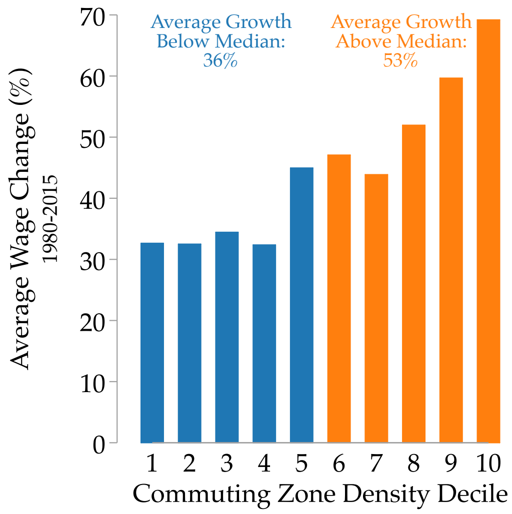
(B) Census/ACS



(C) QCEW

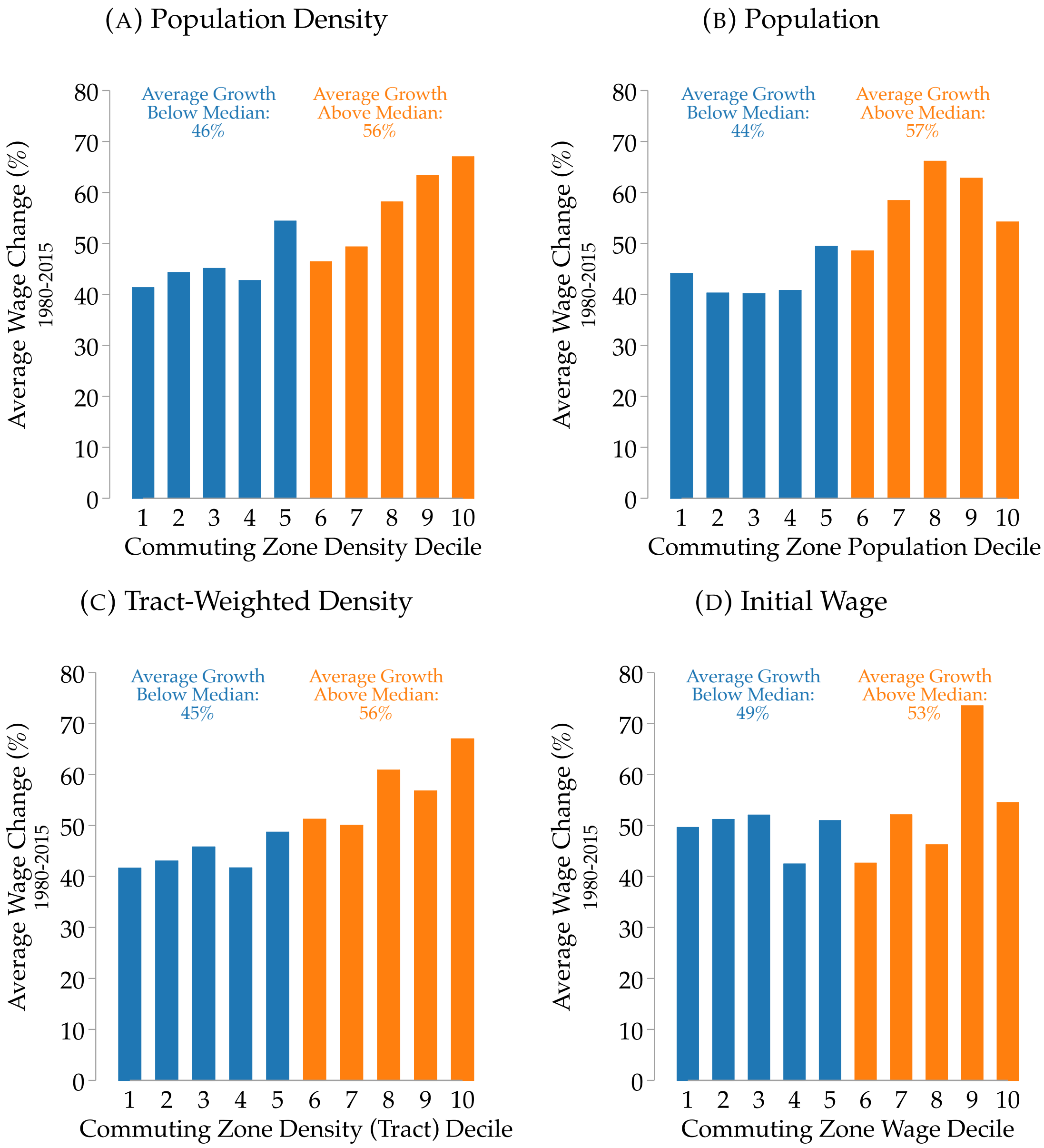


(D) County Business Patterns



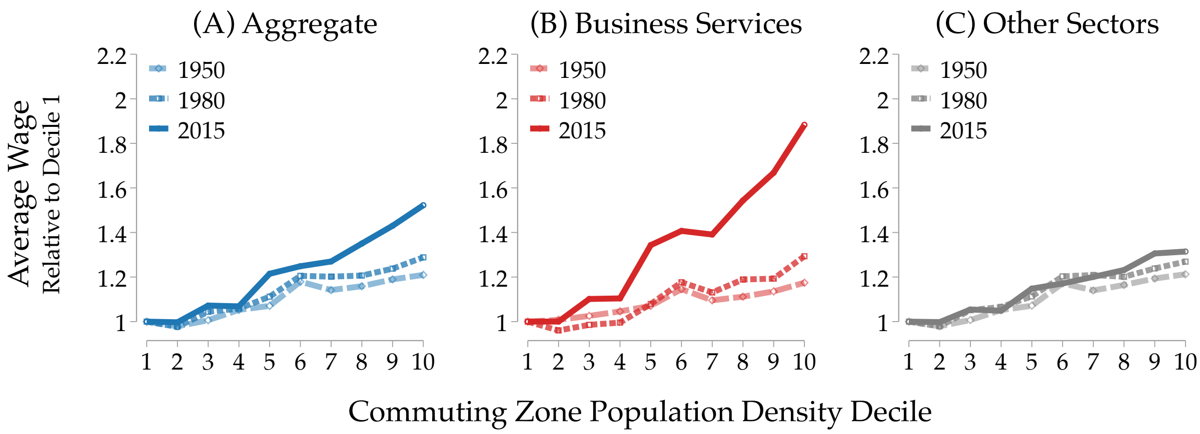
Notes: This figure replicates Figure 2. Panel A replicates using average wages come from the US Census Bureau's LBD. Panel B uses the US Decennial Census/ACS. Panel C uses the BLS QCEW. Panel D uses the Census County Business Patterns files. This figure shows average wages across commuting zones (Tolbert and Sizer, 1996) sorted into 10 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. We define average wages as total payroll over total employment within a commuting zone, deflated by the BEA PCE Deflator.

FIGURE OA.2: WAGE GRADIENT - DIFFERENT ORDERINGS



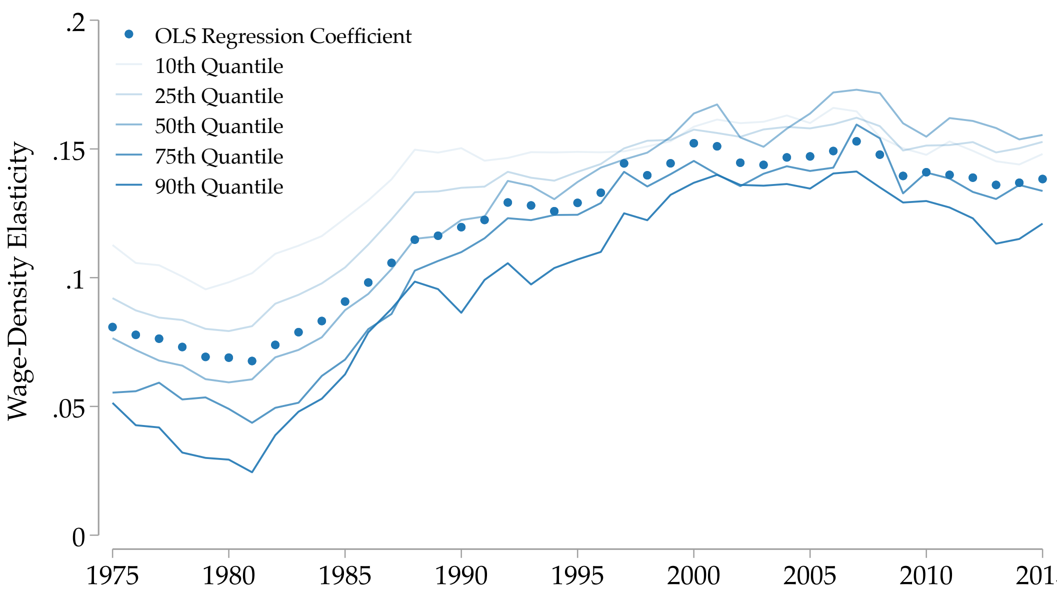
Notes: This figure replicates Panel B of OA.1 using the US Decennial Census/ACS. Panel A replicates the original ordering of commuting zones by initial population density. Panel B orders US commuting zones by initial aggregate population. Panel C uses tract-weighted population density using 1990 data (first year with complete coverage). Panel D orders commuting zones by initial 1980 wage levels.

FIGURE OA.3: Wage Gradient: 1950,1980,2015



Notes: This figure replicates Figure 1 with the addition of 1950 wage gradient data and US Census wage data instead of Census LBD data. This figure shows average wages across commuting zones (Tolbert and Sizer, 1996) sorted into 10 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 $people/mi^2$. Business Services firms are firms in the NAICS-5 sector. We define average wages as total payroll over total employment within a commuting zone and sector pair.

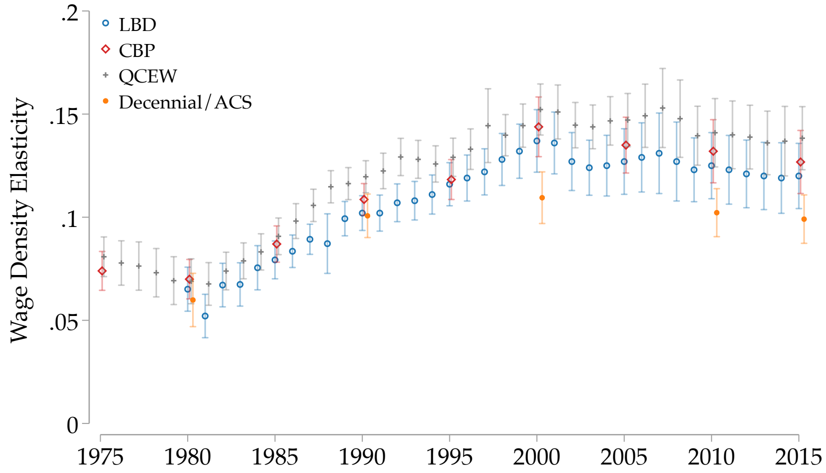
FIGURE OA.4: THE US WAGE-DENSITY GRADIENT OVER TIME



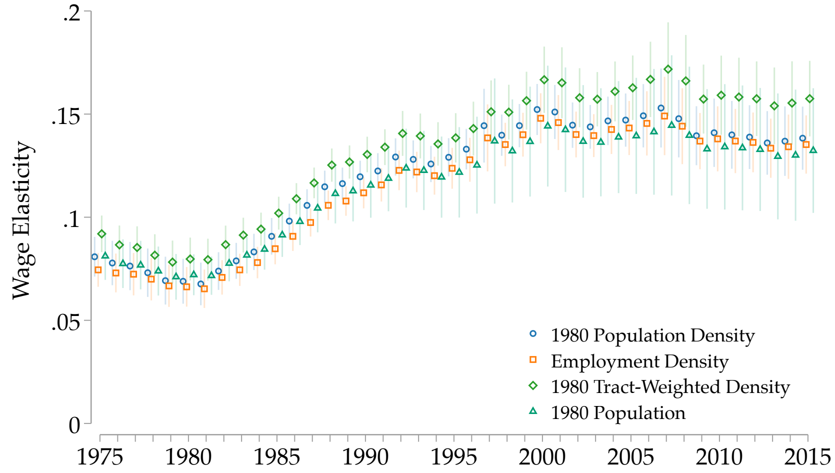
Notes: This figure shows coefficients from a regression of log average wages on log population density run separately for each year between 1975 and 2015 across US commuting zones (blue dots), weighted by 1980 population. We use the US Bureau of Labor Statistics' Quarterly Census of Employment and Wages 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.

FIGURE OA.5: THE US WAGE-DENSITY GRADIENT ROBUSTNESS

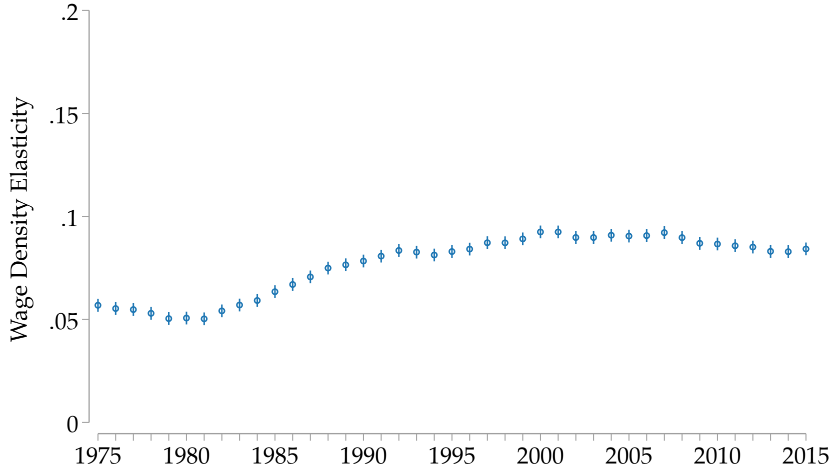
(A) Comparing Data Sources



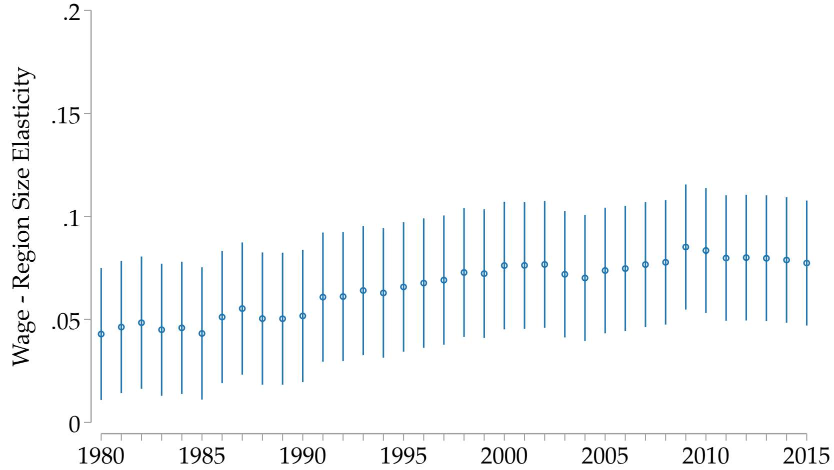
(B) Comparing Density Measures - QCEW



(C) County-Level Data - QCEW



(D) European Union (15) Data



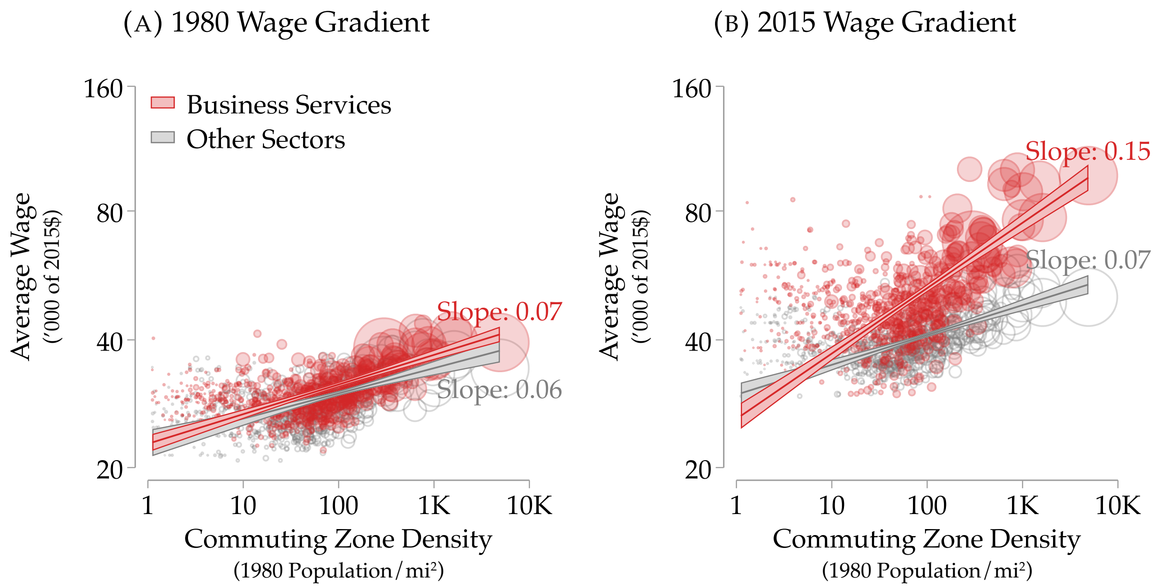
Notes: This figure shows the wage-density gradient coefficients β_t across commuting zones (Panels A and B) and counties (Panel C) for each year from the regression $\ln w_{\ell t} = \alpha + \phi_t + \beta_t \ln density_\ell + \epsilon_{\ell t}$. Panel D replicates Ehrlich and Overman (2020) for all years from 1980-2015. The sample covers EU-15 countries and reports the coefficients β_t across locations, ℓ, for each year from the regression $\ln w_{\ell t} = \alpha + \phi_t + \beta_t \ln employment_\ell + \epsilon_{\ell t}$. $\ln employment_\ell$ refers to the size of the workforce in location ℓ.

TABLE OA.2: THE ROLE OF EDUCATION AND OCCUPATION

	Share of Urban-Biased Growth				
	Education		Occupation		
Sector	Deepening	Residual	Deepening	Residual	Total
Resources + Construction	–0.3	10.1	–1.4	11.2	9.8
Manufacturing	12.4	–36.6	6.8	–31.0	–24.2
Trade + Transport	3.7	–13.8	0.7	–10.8	–10.1
Business Services	16.3	102.0	11.4	106.9	118.3
Education + Medical	1.8	1.0	–1.2	4.0	2.8
Arts + Hospitality	0.7	0.8	–0.1	1.6	1.5
Personal Services	0.2	1.7	0.1	1.8	1.9
Total	35.0	65.0	16.3	83.7	100.0

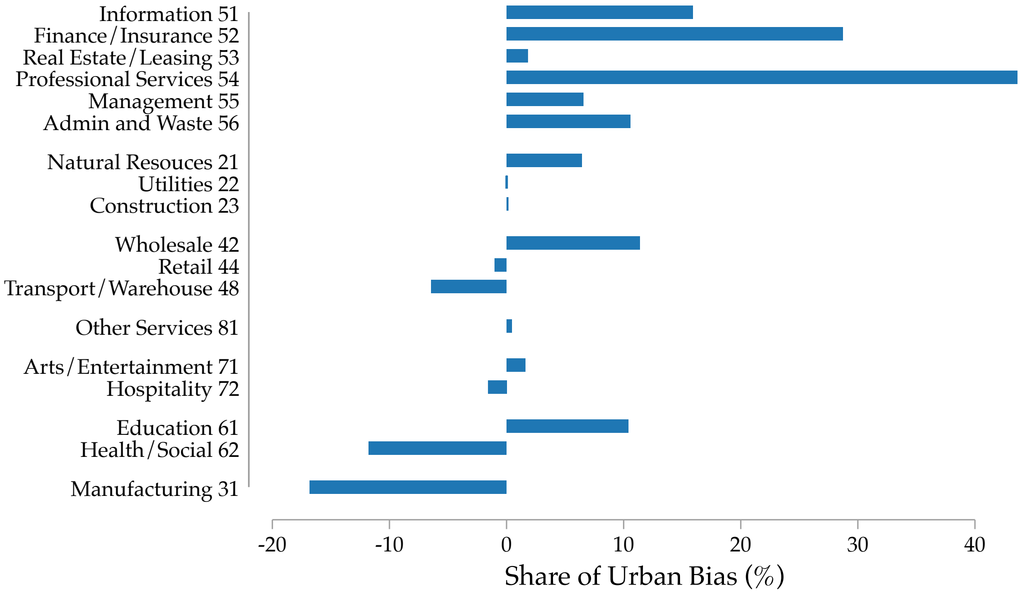
Notes: We use US Census data for 1980 and the ACS data for 2015 from IPUMS and deflate all values by the BEA PCE Deflator. We compute average wages of full-time, prime-age workers within each commuting zone, sector, and either occupation or education group for both years. We follow Jaimovich and Siu (2020), and define CNR occupations with SOC-2 codes 11 to 29 and non-CNR occupations as all other codes. We only consider private non-agricultural employment. Not all observations have an occupational or industry code, and we omit those observations.

FIGURE OA.6: AVERAGE WAGES ACROSS COMMUTING ZONES BY SECTOR IN 1980 AND 2015



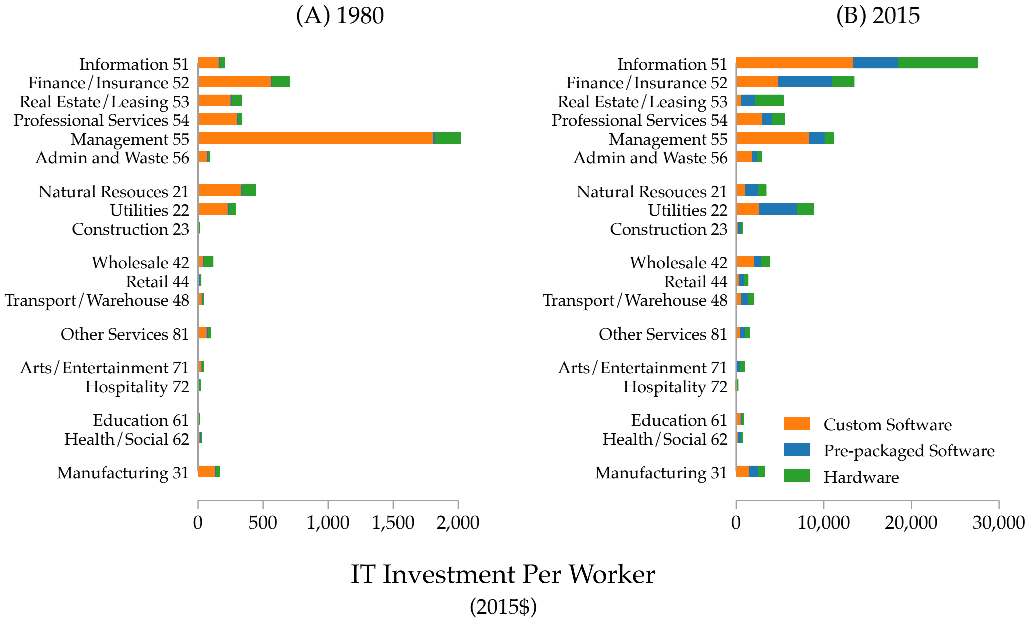
Notes: The left panel of this figure plots average wages at the commuting zone and sector level against commuting zone density in 1980. Size of circles is 1980 population. The right panel does the same for 2015. All wages are in 2015 dollars. The data are from the Decennial Census (1980) and the ACS (2015). The data are adjusted by the BEA PCE Deflator.

FIGURE OA.7: SECTORAL ORIGINS OF URBAN-BIASED WAGE GROWTH ACROSS NAICS-2 INDUSTRIES



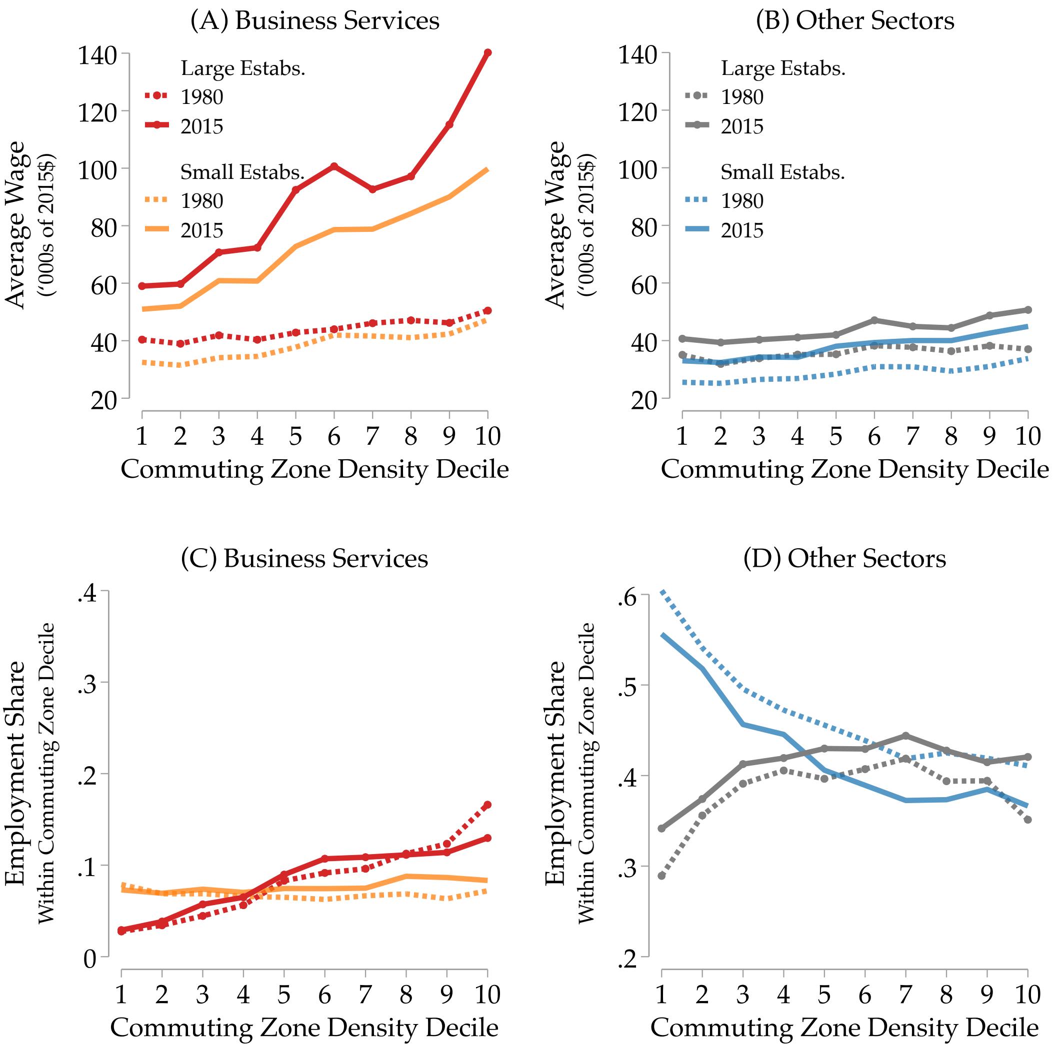
Notes: This figure shows the share of urban-biased wage growth between 1980 and 2015 accounted for by each NAICS-2 industry using the decomposition in equation (2). We compare wage growth between the commuting zones with the highest population density jointly accounting for 50% and all remaining commuting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

FIGURE OA.8: IT INVESTMENT ACROSS NAICS-2 INDUSTRIES



Notes: This figure shows the value of real IT investment per employee by industry in 2015 dollars at the 2-digit NAICS level. Data are from the BEA. Proprietary software refers to BEA codes ENS2 and ENS3, pre-packaged software refers to ENS1, and hardware to EP1A-EP31. Industry groups are ordered by contribution to urban-biased growth.

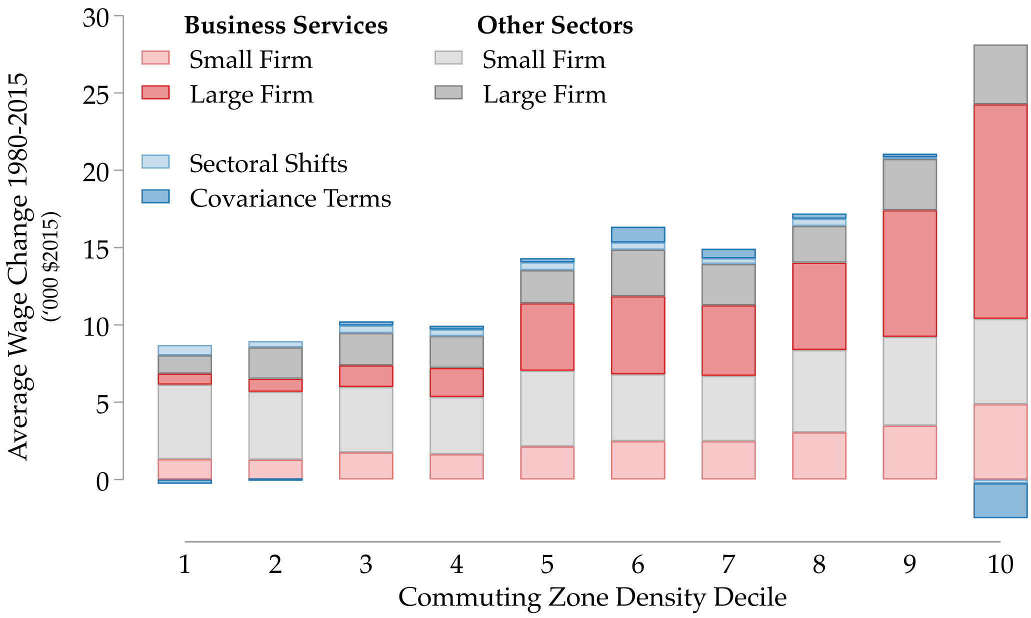
FIGURE OA.9: EMPLOYMENT AND WAGES AT LARGE AND SMALL ESTABLISHMENTS



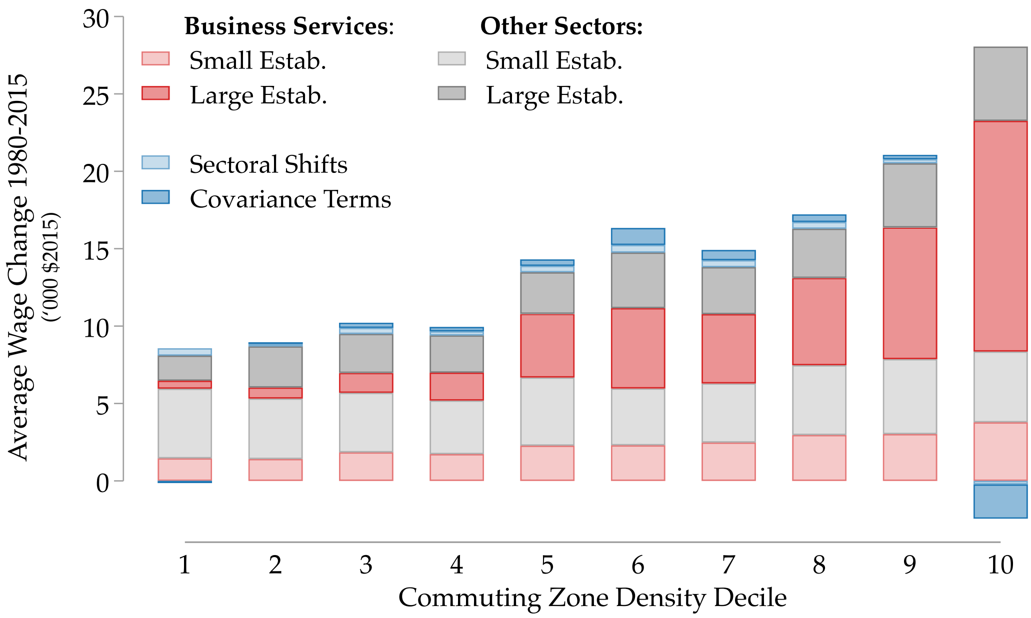
Notes: The top row shows average wages at large and small establishments across commuting zones ordered by population density in 1980 and 2015 for Business Services (Panel A) and all other sectors (Panel B). The bottom row shows employment shares within each commuting zone decile at large and small establishments across commuting zones ordered by population density in 1980 and 2015 for Business Services (Panel C) and all other sectors (Panel D). In all four panels, the solid line reflects 2015, and the dashed line reflects 1980. All wages are stated in 2015 dollars using the BEA PCE Price Index. 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 classify large establishments as those with at least 200 employees, which account for 47% of all employment in 1980 and 50% of all employment in 2015. The data comes from the LBD.

FIGURE OA.10: FIRMS OR ESTABLISHMENTS

(A) The Role of Large Firms

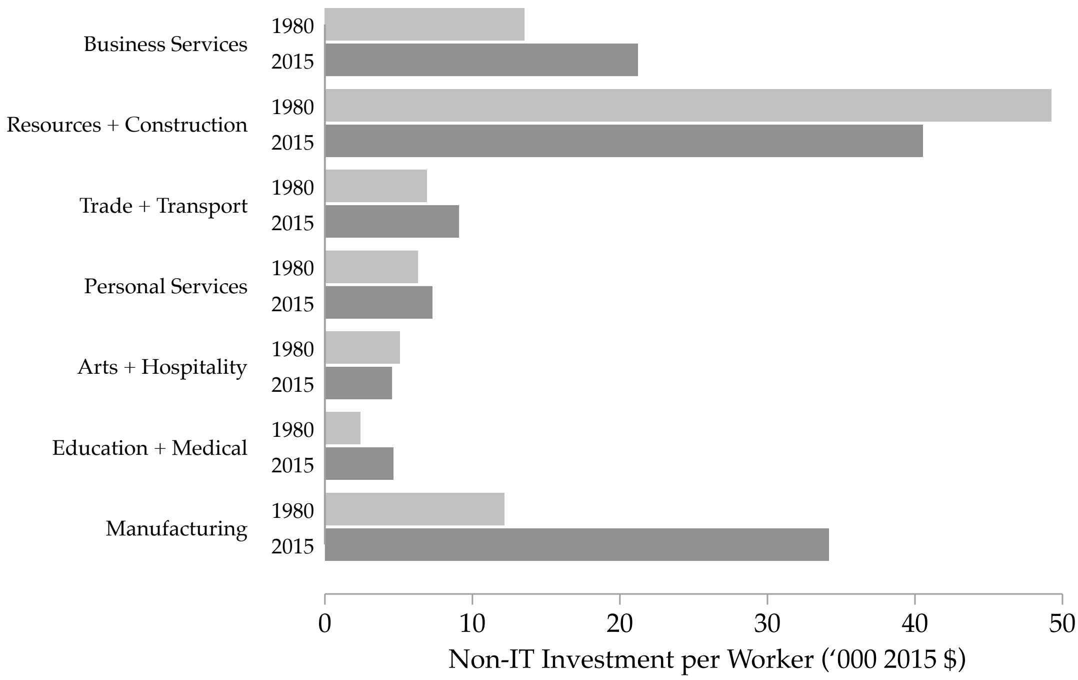


(B) The Role of Large Establishments



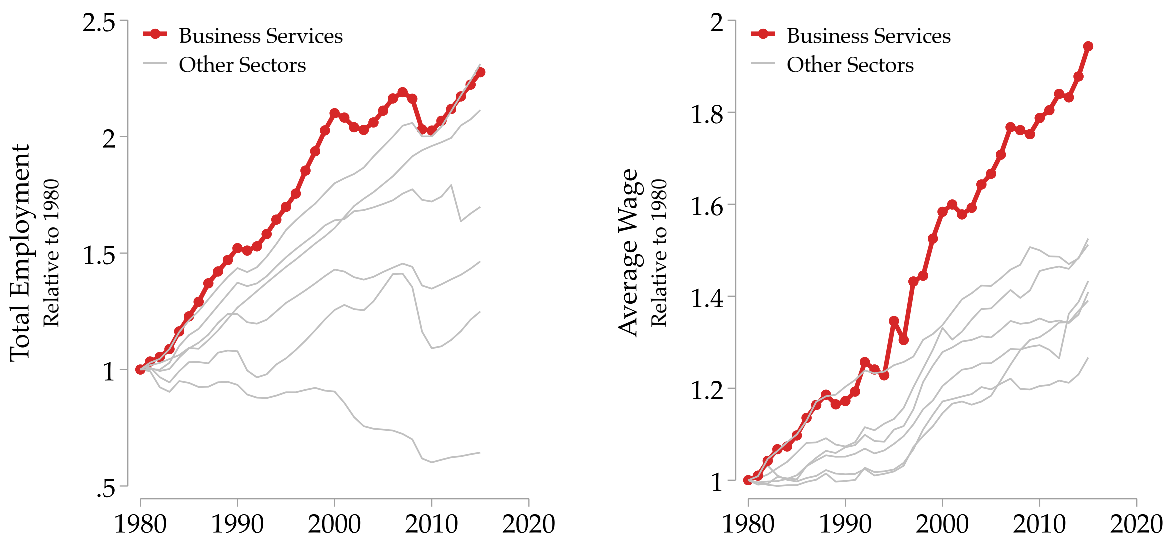
Notes: In the top panel, we classify establishments belonging to large firms as those with at least 1,000 employees in their sector (Business Services versus Other Sectors). Such firms account for roughly 44% of US employment in 1980 and 47% in 2015. To compute the decomposition Business Services firms are those with employment at establishments coded as NAICS 51, 52, 54, 55. Due to disclosure reasons, we omit NAICS 53 and 56 here. In the bottom panel, we classify large establishments as those with at least 200 employees, which account for 47% of all employment in 1980 and 50% of all employment in 2015. We deflate all wages by the BEA PCE Deflator.

Figure OA.11: Aggregate Non-IT Investment by Sector



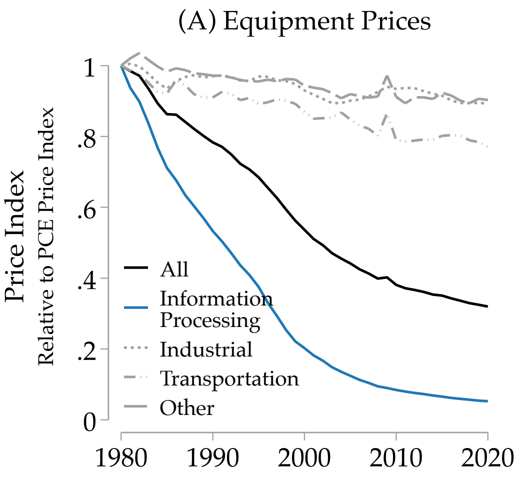
Notes: This figure shows Non-IT Investment Per Worker in 1980 and 2015. Sectors are ordered by their contribution to urban-biased growth. Amounts are deflated using the BEA Fixed-Cost Investment in Private Nonresidential Fixed-Asset tables.

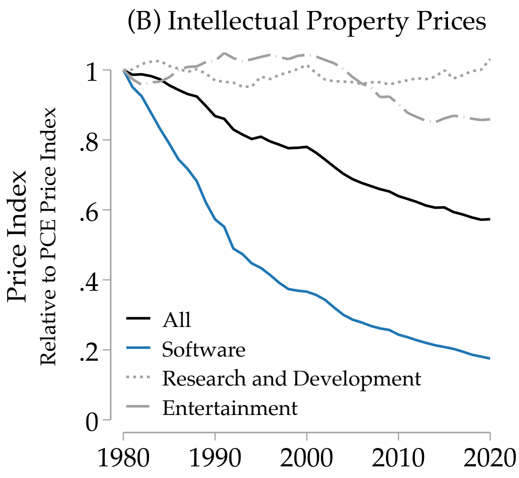
Figure OA.12: Aggregate Wage and Employment Growth



Notes: Panel A shows employment in Business Services (NAICS-5) versus other 1-digit NAICS sectors in the Quarterly Census of Employment and Wages (QCEW). Panel B shows real wages using the same data source deflated by the BEA PCE deflator.

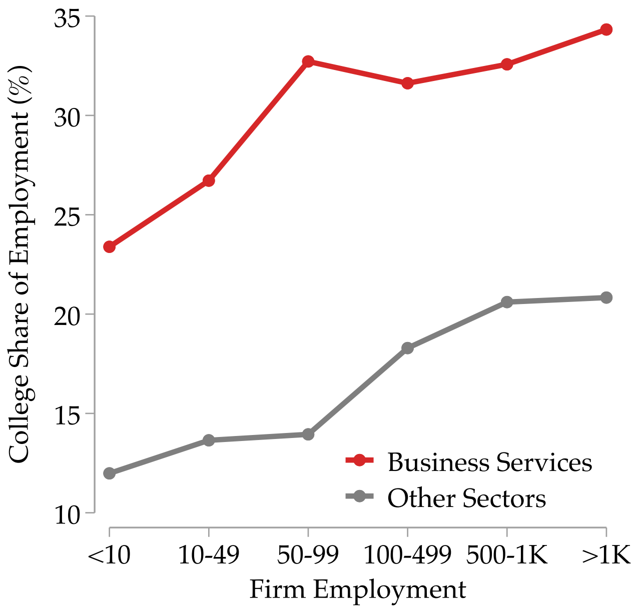
FIGURE OA.13: INVESTMENT PRICE INDICES FOR EQUIPMENT CAPITAL AND INTELLECTUAL PROPERTY





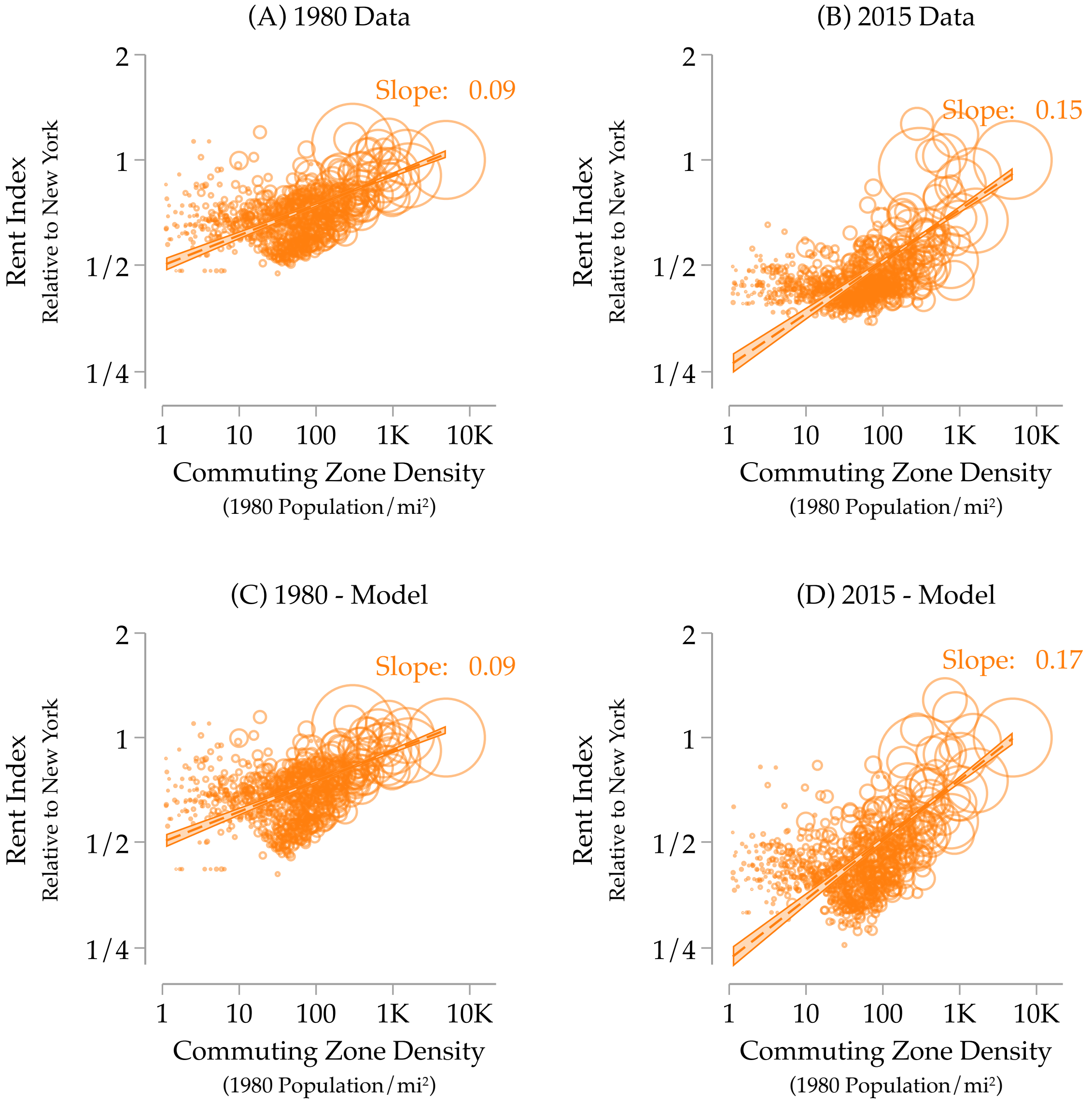
Notes: The left panel of this figure plots the price of equipment investment from the BEA asset price tables for 1980-2018 relative to the BEA PCE deflator. The right panel replicates that plot for intellectual property investment from the BEA asset price tables for 1980-2018 relative to the BEA PCE deflator.

FIGURE OA.14: EDUCATION AND FIRM SIZE



Notes: This figure plots the college educated worker share by firm employment in the 1992 Current Population Survey (CPS) from the US Census.

FIGURE OA.15: COMMUTING ZONE RENT INDEX



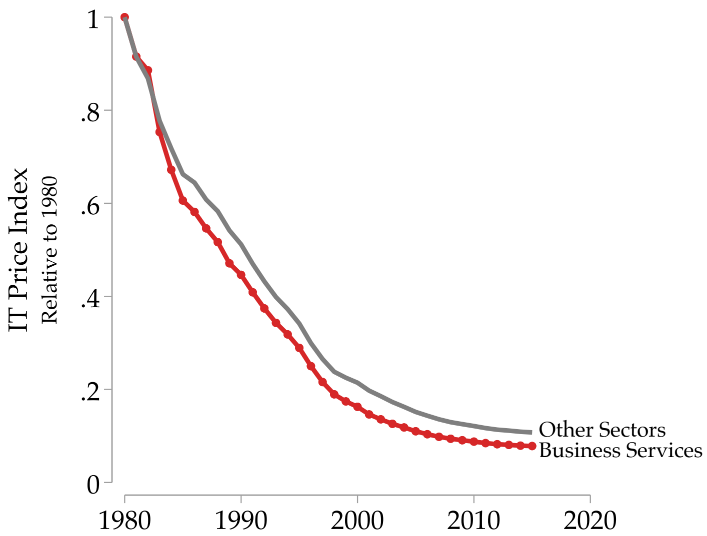
Notes: Panels A and B plot relative rent indexes against commuting zone population density. Mean rent is normalized to 1 in New York. Data are from the US Census and ACS. Panels C and D plot relative rent indexes against commuting zone population density. Rent is normalized to 1 for New York City for that year.

TABLE OA.3: WAGE-DENSITY COEFFICIENT IN DATA AND MODEL: ROBUSTNESS

		2015					
	1980	Data	Only A	Only B	High Elast.	Low Elast.	Equal Lab. Elast.
Business Services	0.070	0.154	-0.027	0.134	0.212	0.100	0.147
Other Sectors	0.060	0.070	0.054	0.059	0.073	0.066	0.068
Aggregate	0.063	0.102	0.042	0.082	0.147	0.078	0.099
Δ Aggregate		0.039	-0.021	0.018	0.084	0.015	0.038

Notes: Gradients computed use the ACS/Census for 1980 and 2015 under different specifications of the model, weighting by 1980 population shares.

FIGURE OA.16: The Decline of the IT Investment Price Index by Sector



Notes: This figure computes the Idea Chained IT Cost/Price index Relative to PCE Chained Price Index, 1970=1. We compute a sectoral ideal (Fischer) price index taking the geometric average of the Laspeyres and Paasche price indices. We do so separately by sector because different sectors have different weights for various equipment types. These two time series use the BEA Asset Price data.

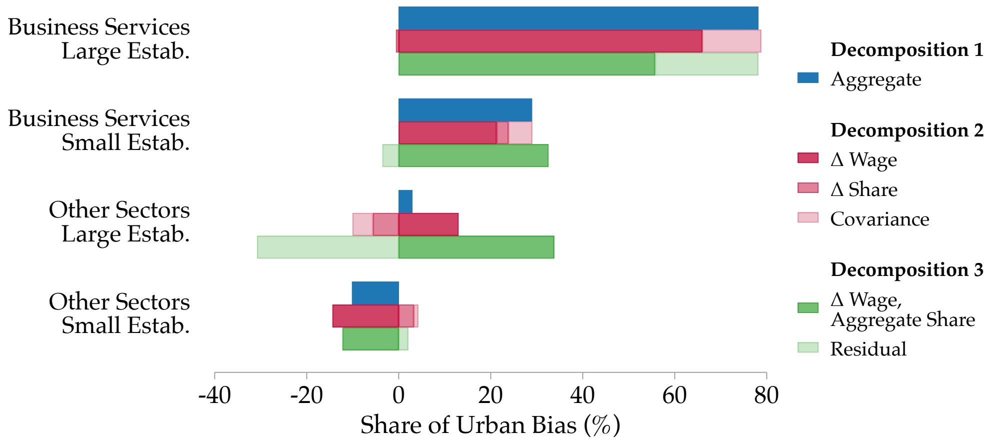
TABLE OA.4: IT EXPENDITURE, POPULATION DENSITY, AND ESTABLISHMENT SIZE IN THE SPICEWORKS DATA

	IT Expenditure/Worker (x $1,000)					
	(1)	(2)	(3)	(4)	(5)	(6)
Log Population Density	0.237***	-0.0222**			0.232***	0.132***
	(0.0124)	(0.00847)			(0.0242)	(0.0295)
Log Employment			0.356***	0.288***	0.317***	0.454***
			(0.00914)	(0.0124)	(0.0344)	(0.0343)
Log Density × Log Emp.					0.00878	-0.0319***
					(0.00728)	(0.00722)
Business Services		0.368*		2.774***		1.334***
		(0.165)		(0.0679)		(0.239)
× Log Density		0.551***				0.252***
		(0.0337)				(0.0486)
× Log Emp.				0.144***		-0.243**
				(0.0182)		(0.0787)
× Log Density × Log Emp.						0.0739***
						(0.0165)
Establishments	2,872,954	2,872,954	2,872,954	2,872,954	2,872,954	2,872,954
R^2	0.000	0.005	0.002	0.007	0.002	0.007

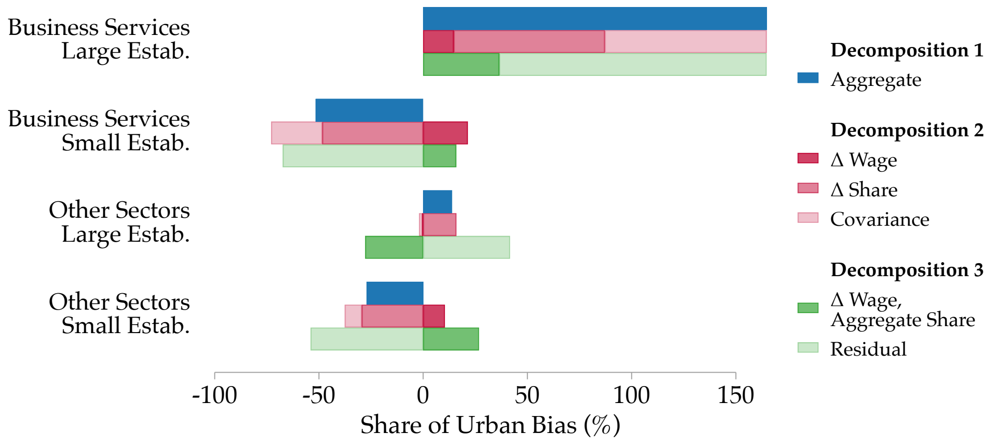
Notes: Standard errors in parentheses. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. The independent variable ”IT expenditure per worker” is in thousands of 2015 dollars. The Business Service indicator represent a Business Service (NAICS-5) establishment. Density is the 1980 population density of the commuting zone. Emp reflects total employees reported. The data come from 2015 Harte-Hanks Market Intelligence.

Figure OA.17: Model and Data: Firm Decomposition

(A) Data



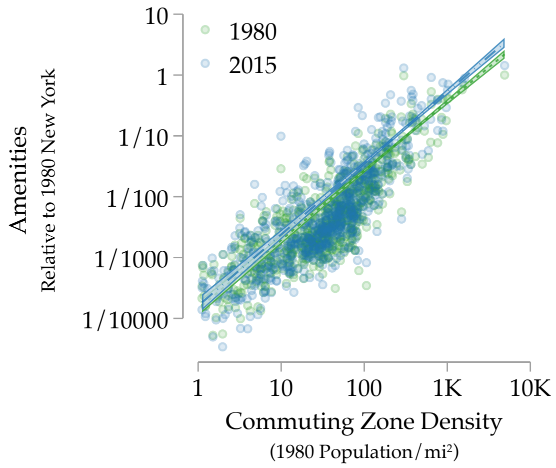
(B) Model



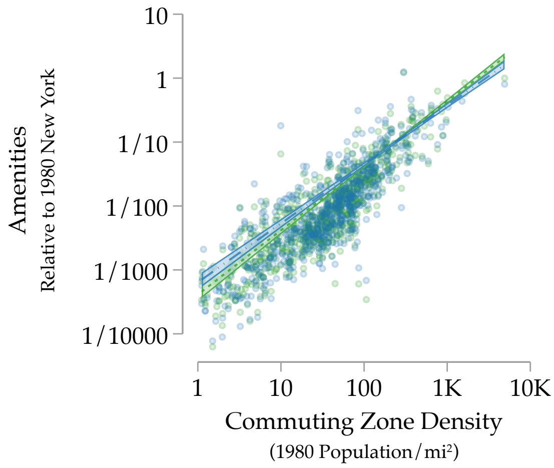
Notes: This figures recreates Figure 4 in the data (Panel A) and the model (Panel B).

Figure OA.18: Amenities in the Model

(A) College



(B) Non-College



Notes: This figure plots relative amenities indexes for 1980 against commuting zone population density. The 2015 amenities are only for the robustness exercise with endogenous amenities. Amenities are normalized to 1 for New York in 1980 for each education group.