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

A. ADDITIONAL FIGURES AND TABLES

A.1 Details on the Urban Wage Gradient Over Time

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

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

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

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

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

A.2 College to Non-College Ratio by Sector

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

A.3 Disaggregated Industry Detail within Sectors

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

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

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

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

(OA.1)

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

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

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

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

A.4 Moments Used in the Model Calibration

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

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

A.5 Additional Model Results

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

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

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

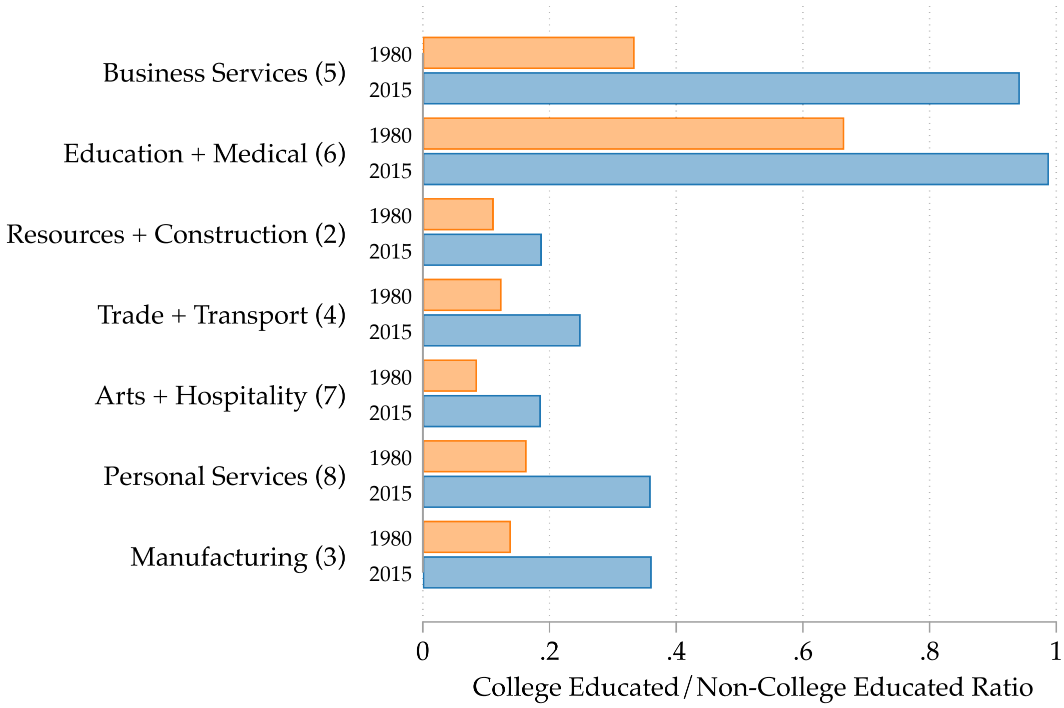
Figure OA.10 plots the fit of the model for ICT per employee and the college share of employment across different firm size bins against data from the ACES (Panel (a)) and the CPS (Panel (b)).

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

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

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

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



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

A.6 Endogenous Local Fundamentals

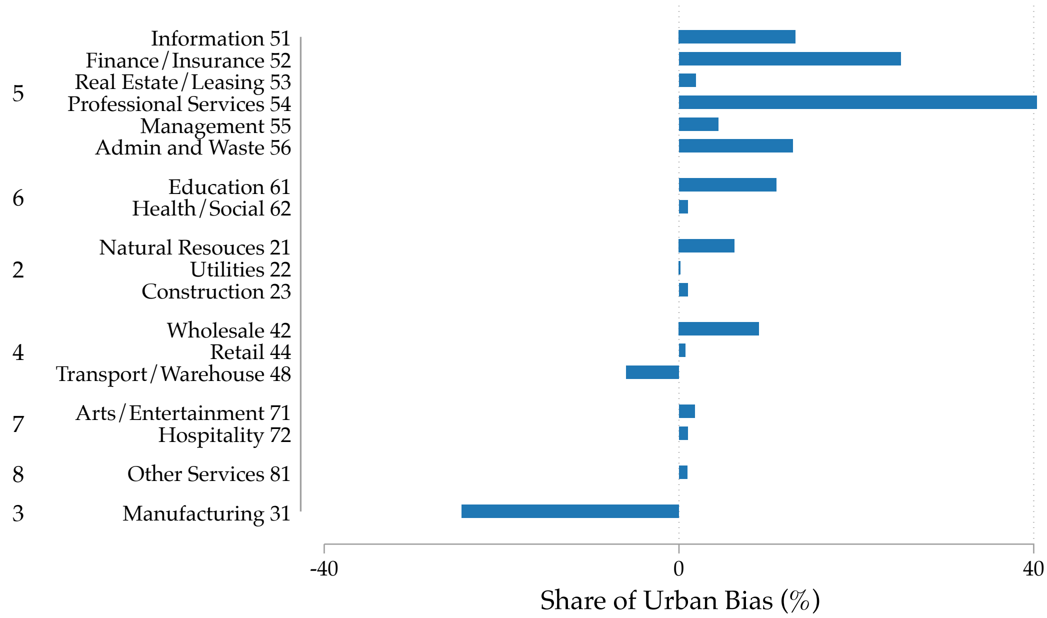
A long literature suggests that local productivities and amenities may be endogenous functions of the size and composition of a location's workforce. In our main calibration, we abstracted from such "spillover" effects. We investigate their qualitative role in affecting the strength of our mechanism.

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

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

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

Figure OA.2: Sectoral Origins of Urban-Biased Wage Growth across NAICS-2 industries



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

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

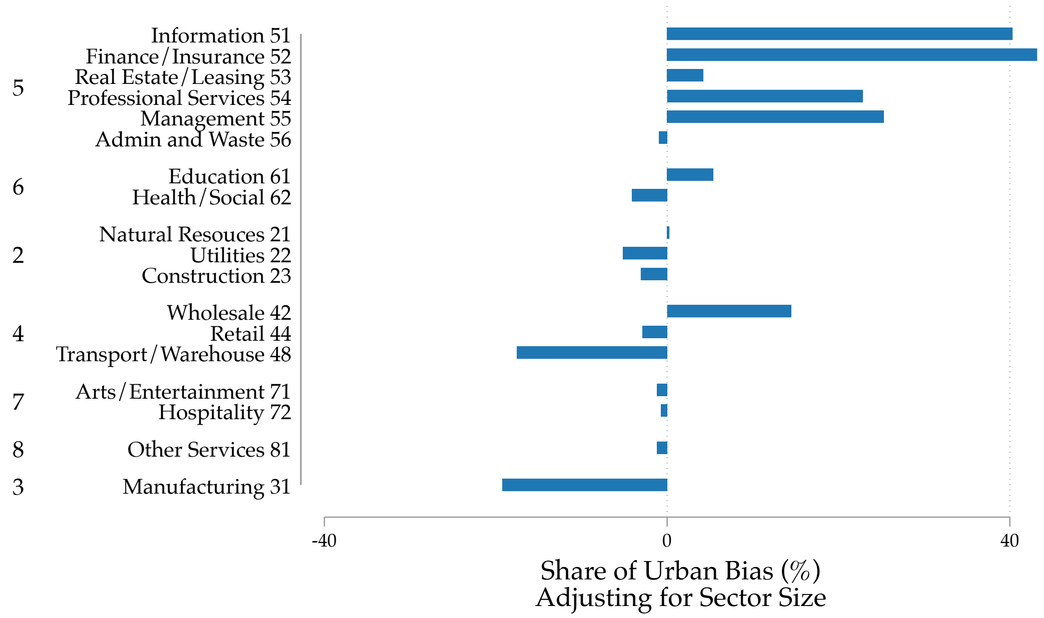
Table OA.1: Location Fundamentals and Employment Density

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

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

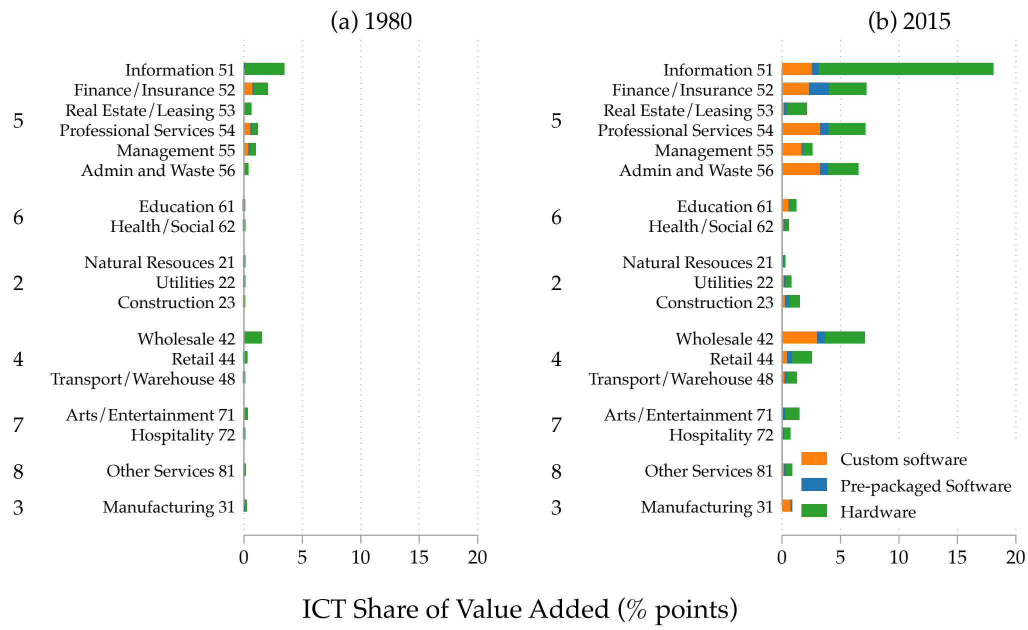
Notes: This table presents six regressions of calibrated model objects for 1980 on commuting zone density. The top panel regresses (log) density on the underlying productivity ($\log Z$) across commuting zones r for four different groups of workers, those in Business Services (N5) and those in other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

FIGURE OA.3: SECTORAL ORIGINS OF URBAN-BIASED WAGE GROWTH ACROSS NAICS-2 INDUSTRIES WITH INDUSTRY SIZE ADJUSTMENT



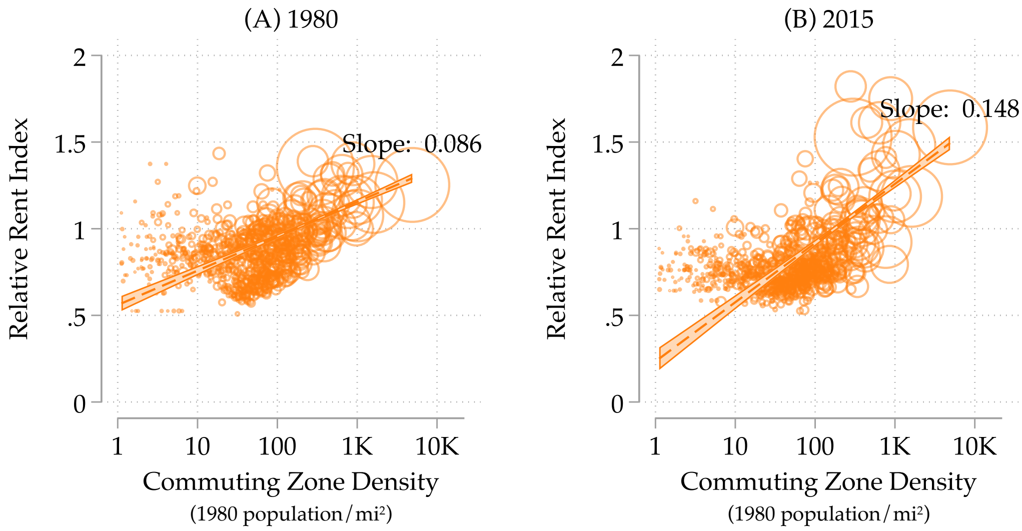
Notes: This figure shows the share of urban-biased wage growth between 1980 and 2015 accounted for by each NAICS-2 industry using the sizeless growth decomposition shown in equation OA.1. We compare wage growth between the commuting zones with the highest population density jointly accounting for 50%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

FIGURE OA.4: ICT VALUE ADDED SHARES ACROSS NAICS-2 INDUSTRIES



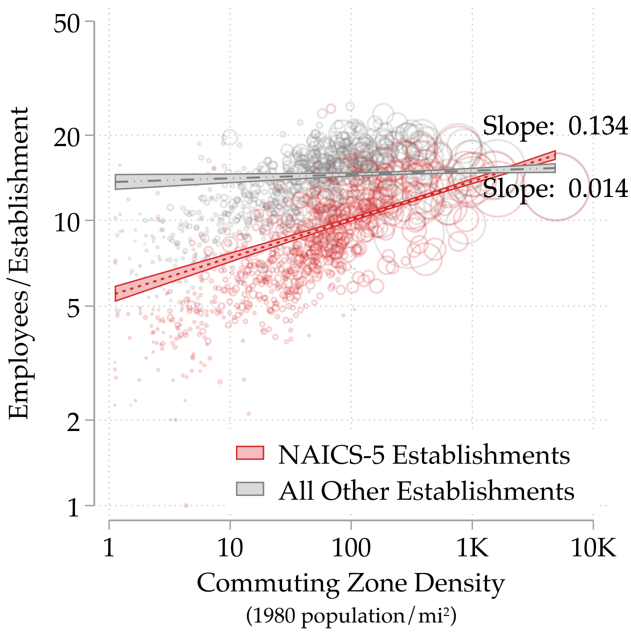
Notes: We show the share of real ICT of value added by industry in 2012 dollars from equation (12). Data are from the BEA. Prior to 1987, labor share uses data from the QCEW. Proprietary software refers to BEA codes ENS2 and ENS3, pre-packaged software refers to ENS1, and hardware to EP1A-EP31. Sectors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

FIGURE OA.5: COMMUTING ZONE RENT INDEX



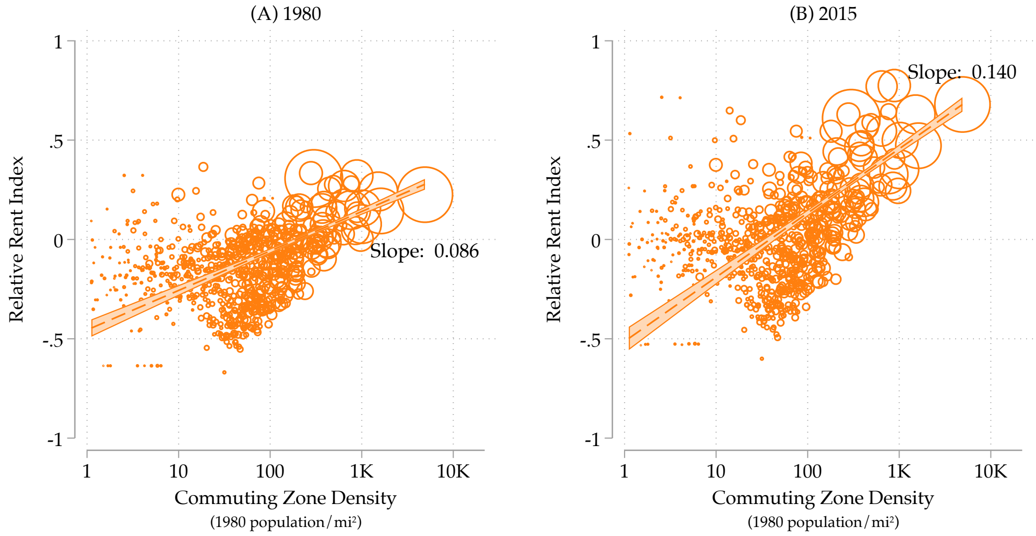
Notes: This figure plots relative rent indexes against commuting zone population density. Mean rent is normalized to 1. Data are from the US Census and American Community Survey.

FIGURE OA.6: AVERAGE ESTABLISHMENT SIZE ACROSS SPACE



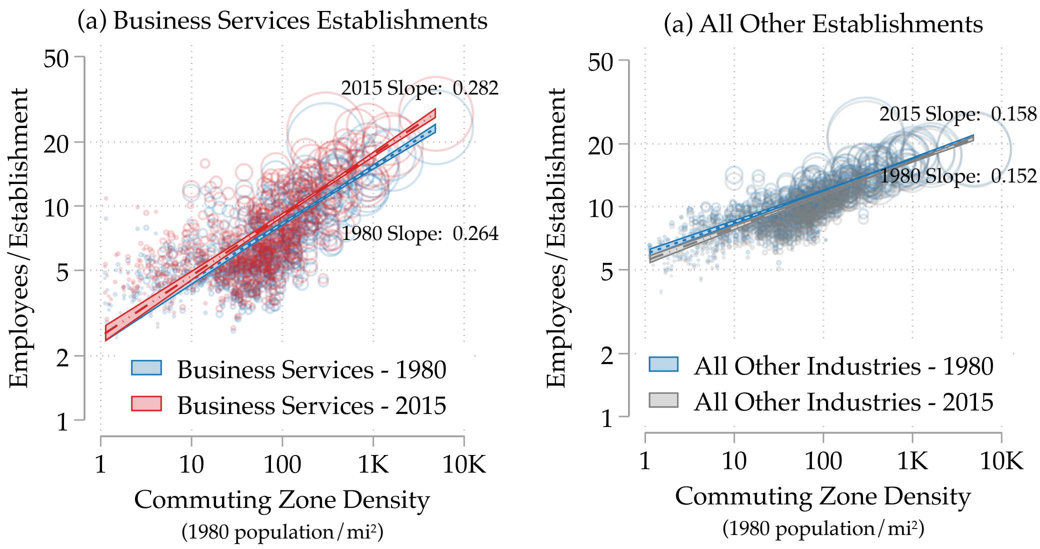
Notes: This figure plots average establishment size against population density using data from the 2015 QCEW. The regression coefficient is weighted by commuting zone population.

FIGURE OA.7: COMMUTING ZONE RENT INDEX IN THE MODEL



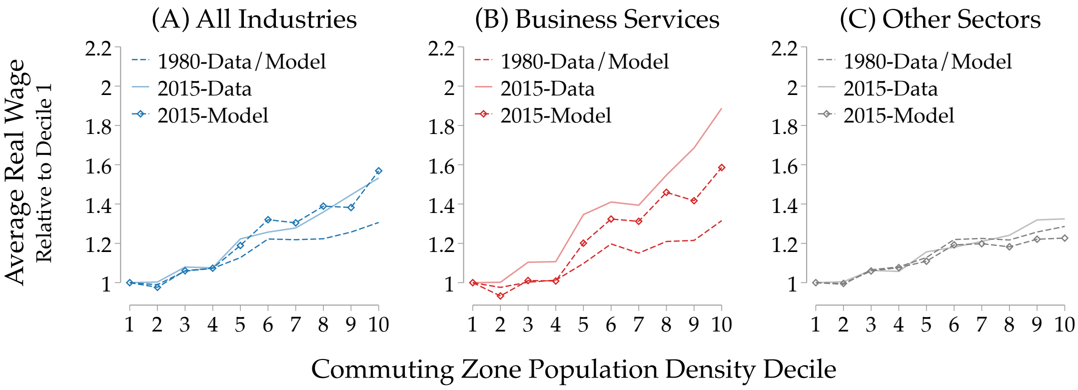
Notes: This figure plots relative rent indexes against commuting zone population density. Mean rent is normalized to 0.

FIGURE OA.8: AVERAGE ESTABLISHMENT SIZE ACROSS SPACE IN THE MODEL



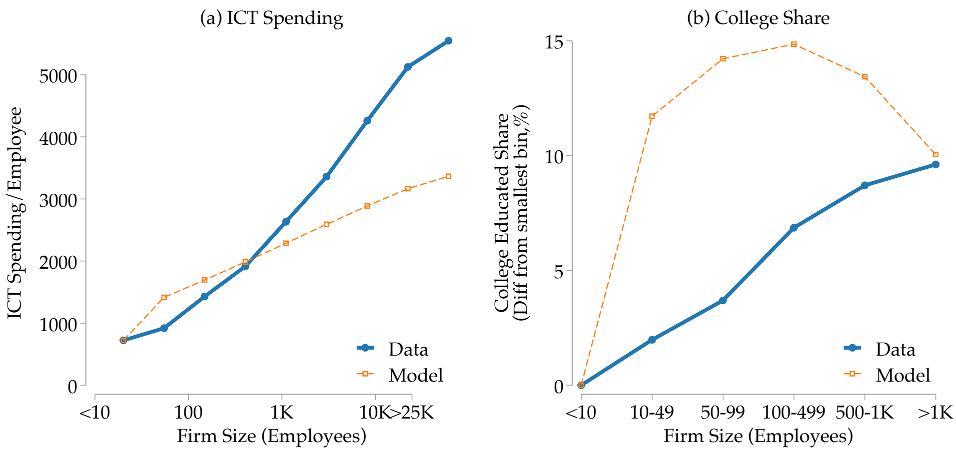
Notes: This figure plots average establishment size against population density in the calibrated model for both 1980 and 2015.

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



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

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



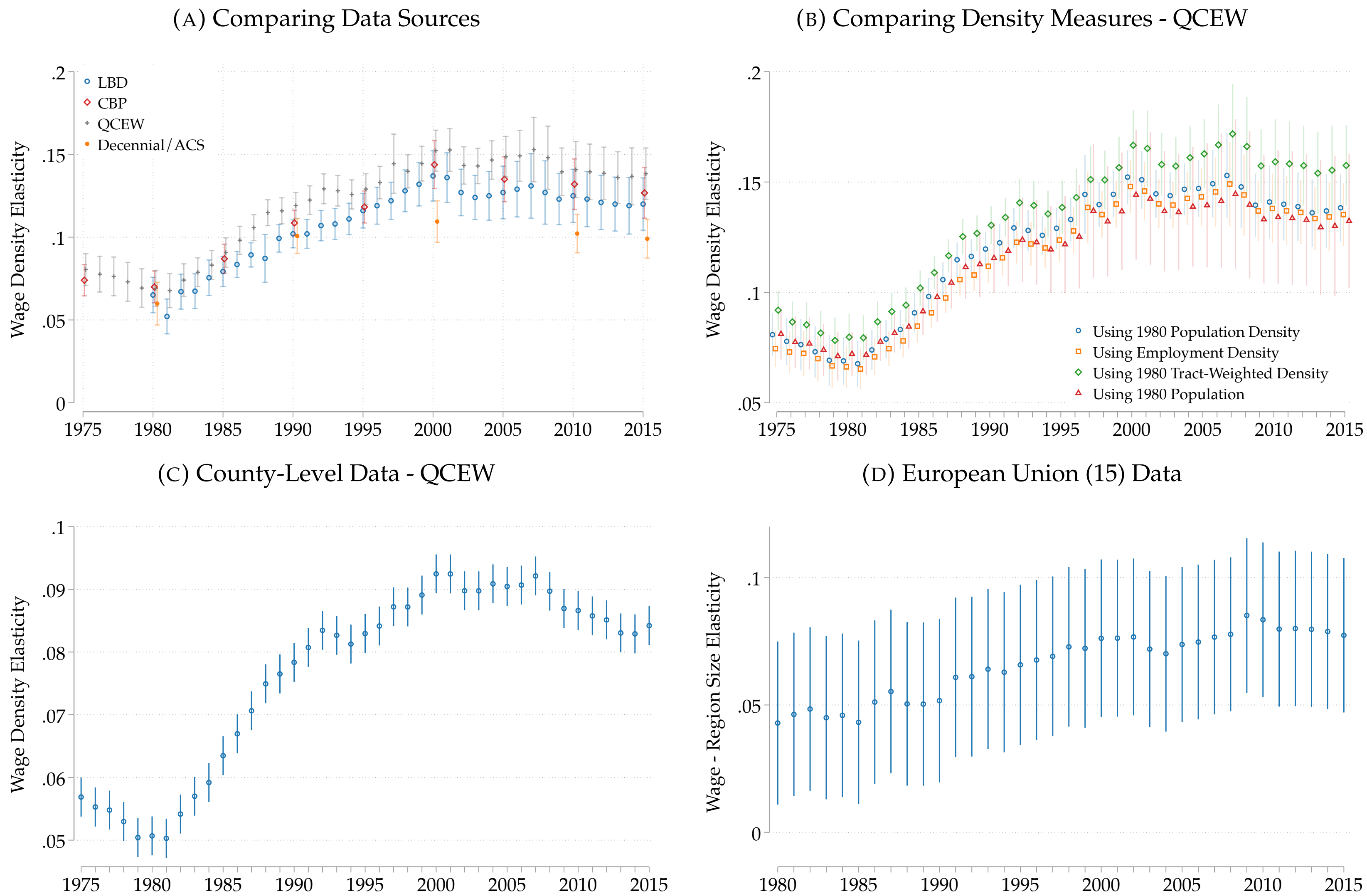
Notes: The left panel plots ICT spending per worker using the 2013 ACES/ICT survey matched to the LBD from the US Census. The right panel plots the difference in college educated worker share compared with the firms with under 10 employees in the 1992 Current Population Survey (CPS) from the US Census.

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

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

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

FIGURE OA.11: WAGE-DENSITY COEFFICIENTS



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

B. Theory Appendix

B.1 The Baseline Model

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

B.1.1 The Firm's Problem

An individual firm produces with the following production technology:

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

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

We can write the firm's cost minimization problem:

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

The resulting first order conditions are given by:

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

Plugging the first order conditions back into the cost function:

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

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

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

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

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

which satisfies

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

Thus we have

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

Now from the definition of profit

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

From (OA.8) we have

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

and so we can write

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

Similarly for the effect of productivity on profit

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

B.1.2 Useful Lemmas

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

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

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

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

or

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

□

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

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

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

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

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

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

□

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

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

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

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

But then:

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

□

B.1.3 Proof of Proposition 1

We can totally differentiate the free entry condition to find

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

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

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

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

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

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

and plugging this into Lemma 3

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

Then using (OA.11) we get

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

or

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

Note that the denominator is always positive since

$$1+\Theta_r\epsilon-\zeta>1>\epsilon^2\theta_r$$
$$(1-\Theta_r)(1-\sigma)<1<(1+\theta_r\epsilon).$$

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

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

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

B.1.4 Proof of Proposition 2

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

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

which can be written

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

and substituting in the free entry condition yields

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

As a result we have

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

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

B.1.5 Changes in Aggregate Demand

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

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

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

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

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

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

which we can re-arrange to yield:

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

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

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

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

B.2 The Neoclassical Baseline

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

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

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

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

$$L_r = M_r(w_r)$$

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

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

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

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

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

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

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

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

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

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

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

Using the definition of the cost shares:

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

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

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

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

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

Now totally differentiate to obtain:

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

Finally, re-arranging:

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

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

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

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

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

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

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

C. Data Construction

C.1 Defining Commuting Zones

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

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

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

C.2 Price Index Data

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

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

to 2015.

C.3 LBD

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

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

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

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

C.4 Census

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

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

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

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

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

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

C.5 QCEW

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

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

C.6 CPS

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

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

C.7 CBP

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

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

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

C.8 BEA Fixed Asset and Value Added Data

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

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

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