SUPPLEMENTARY MATERIAL FOR

URBAN-BIASED GROWTH: A MACROECONOMIC ANALYSIS

BY FABIAN ECKERT, SHARAT GANAPATI, AND CONOR WALSH

September 2022

NOT FOR PUBLICATION

A. Industries, Occupations or Education

A large recent literature focuses on the role of occupations in employment polarization in the United States. Papers in this literature often categorize occupations by their task content. We follow Rossi-Hansberg et al. (2019) and Rossi-Hansberg et al. (2019) in classifying occupations into cognitive non-routine occupations (CNR) and others (non-CNR). CNR occupations are typically high-skill occupations that require cognitive non-routine abilities. The left panel of Figure SM.1 shows wage growth across US commuting zones ordered by their density in 1980, separately for workers in CNR and non-CNR occupations. CNR wages did grow faster in denser locations. But the figure also reveals that CNR and non-CNR occupations *not* in the business services sector exhibit wage growth that is largely unbiased across space. On the other hand all occupations within the business services sector experienced wage growth strongly biased toward denser labor markets. The figures suggests that the density-bias in CNR wage growth is driven by the fact that business service industries employ a disproportionate amount of CNR workers, but is not particular to CNR occupations.

Similarly, the bias in recent wage growth toward dense location may reflect something about education. Recent papers (e.g., Giannone (2022) and Rubinton (2019)) argue that the disproportionately fast growth of skilled wages in many large urban areas is due to faster skill-biased technical change in such cities. The right panel of Figure SM.1 shows that indeed growth in wages for college-educated workers has exhibited a stronger density bias than growth in non-college wages. However, once we condition on the Business Services sector we see that this is driven almost entirely by compositional differences: both college and non-college wages in business service exhibit a strong bias toward denser labor markets, while not much of a bias exists for wages of these groups outside business services. The reason skilled wages have seen faster growth in denser areas is their over-representation in the Business Services sector, it is however not specific to skilled workers per se.[46]

We use data from the Decennial Census and ACS to look at wage differences within occupations (defined at the 6-digit SOC code level). For every commuting zone, educational group, year, business service classification, 6-digit

[46]Figure SM.1 also shows the disappearance of the low-skill urban wage premium: non-college workers experienced negative wage growth in some of the densest local labor markets (deciles 6,7, and 8).

occupational code, we regress real wages on a business service dummy and an business service dummy interacted with CZ density, controlling for a full set of interactions between, occupations, educational groups, and CZ. We weight the regressions by the number of employees within each cell. The results are presented in Table SM.1. Column (1) shows the results without interacted fixed effects and Column (2) interacts all the fixed effects.[47] In 1980, there was no aggregate wage premium within occupation codes for business services. By 2015, the wage premium ranged between $4,000 and $7,000 within a the same city, educational, group, and 6-digit SOC occupational code. Column (3) interacts this wage premium with CZ density in 1980 (scaled between 0 and 1). In particular, we find that the entirety of the business service wage premium within jobs is driven by the densest CZ, with no difference for the least dense CZ. Further analysis reveals wide variation in wages driven by the large number of employees as "Managers and administrators," "Chief executives and public administrators," "Accountants and auditors," and "Office supervisors."

In summary, the evidence in this section suggests that the recent changes in the economic geography of wage growth are indeed driven by the Business Services sector, rather than certain skill or occupational groups.

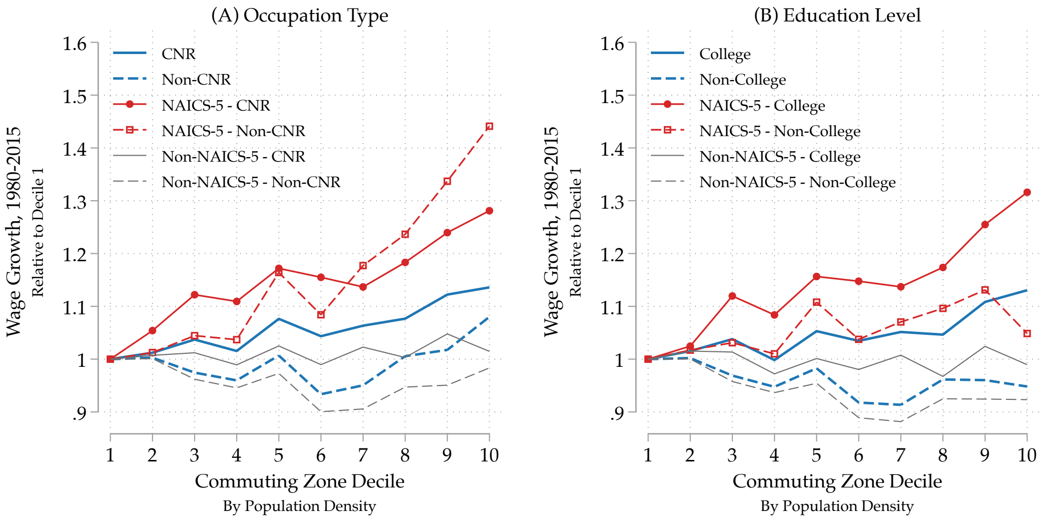
[47]The number of observations decreases due to the number of singleton cells.

TABLE SM.1: REAL WAGES DIFFERENCE WITHIN OCCUPATION

	(1) Real Wage	(2) Real Wage	(3) Real Wage
1980 × NAICS-5 Premium	-2867.0	-800.6	-1545.0
	(1168.7)	(424.0)	(542.7)
1990 × NAICS-5 Premium	452.1	616.2	-2213.7
	(763.1)	(608.8)	(576.7)
2000 × NAICS-5 Premium	3482.3	1714.1	-2431.5
	(972.8)	(602.7)	(466.0)
2010 × NAICS-5 Premium	4119.8	3178.7	-2474.7
	(994.0)	(903.9)	(618.3)
2015 × NAICS-5 Premium	6188.8	4283.0	-1795.6
	(1115.9)	(920.0)	(727.9)
1980 × NAICS-5 Premium × CZ Density Decile			1160.2
			(405.7)
1990 × NAICS-5 Premium × CZ Density Decile			4475.7
			(390.1)
2000 × NAICS-5 Premium × CZ Density Decile			6743.3
			(817.2)
2010 × NAICS-5 Premium × CZ Density Decile			9082.3
			(1394.3)
2015 × NAICS-5 Premium × CZ Density Decile			9795.4
			(1469.7)
Constant	44839.6	47211.5	47192.9
	(124.7)	(137.0)	(135.7)
Observations	3695157	1489900	1489900
R^2	0.714	0.930	0.930
FE: Yr, Occupation, Education, CZ	✓		
FE: Yr x Occupation x Education x CZ		✓	✓

Notes: Robust standard errors in parentheses. Observations are at a sector-commuting zone-6 digit SOC occupational code- educational (HS or less, Some College, College, or Graduate Degree)-commuting zone cell. Commuting zone density decile scaled from 0 to 1, with 0 representing the least dense commuting zones. The data are from the Decennial Census and the ACS, adjusted by the BLS CPI-U.

FIGURE SM.1: AVERAGE WAGE GROWTH ACROSS COMMUTING ZONES BY SECTOR, OCCUPATION, AND EDUCATION GROUP



Notes: This figure plots average wage growth by occupation (Panel (A)) and education (Panel (B)) across the 10 density decile groupings used in the paper, relative to the first decile. The data are from the Decennial Census (1980) and the ACS (2015), adjusted by the BLS CPI-U.

B. Additional Figures

B.1 Wage Convergence and Divergence - Additional Figures

Figure SM.2 shows commuting zone level wage convergence in the decade from 1970-1980, while SM.3 shows partial divergence in the period under study in the paper. Both show wage growth as a function of initial wage level.

Figure SM.4a shows wage growth as a function of population density, where the role of large dense cities is more apparent.

Figure SM.4b show the results of the divergent growth are primarily driven by wages with the business services sector (NAICS-5), with the coefficient of a regression of wage growth on population density over five times higher than for the rest of the economy.

Figure SM.4c shows the result of this period of divergent growth, with the coefficient of a regression of wages on population density almost doubling in the period under study.

FIGURE SM.2: WAGE CONVERGENCE, 1970-1980

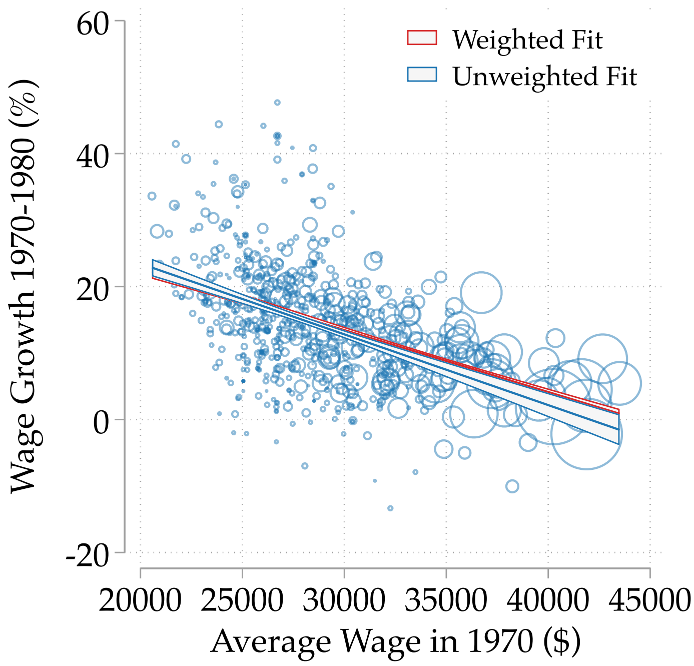
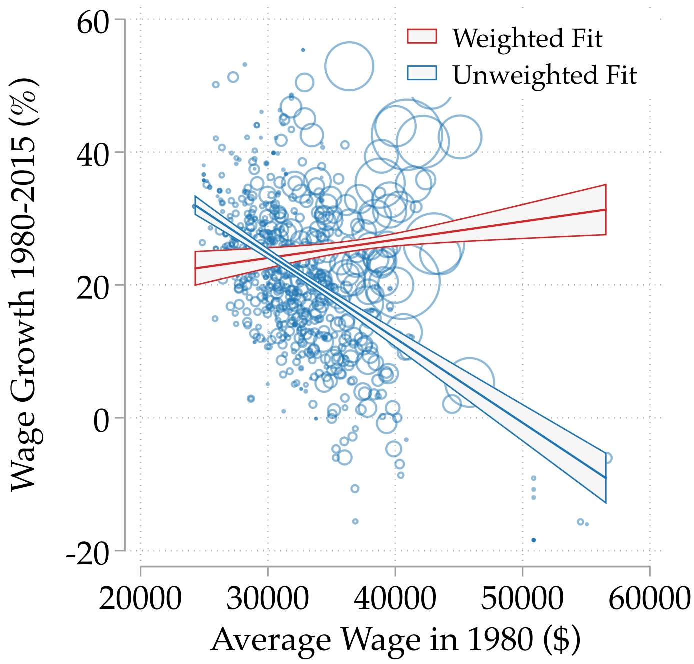
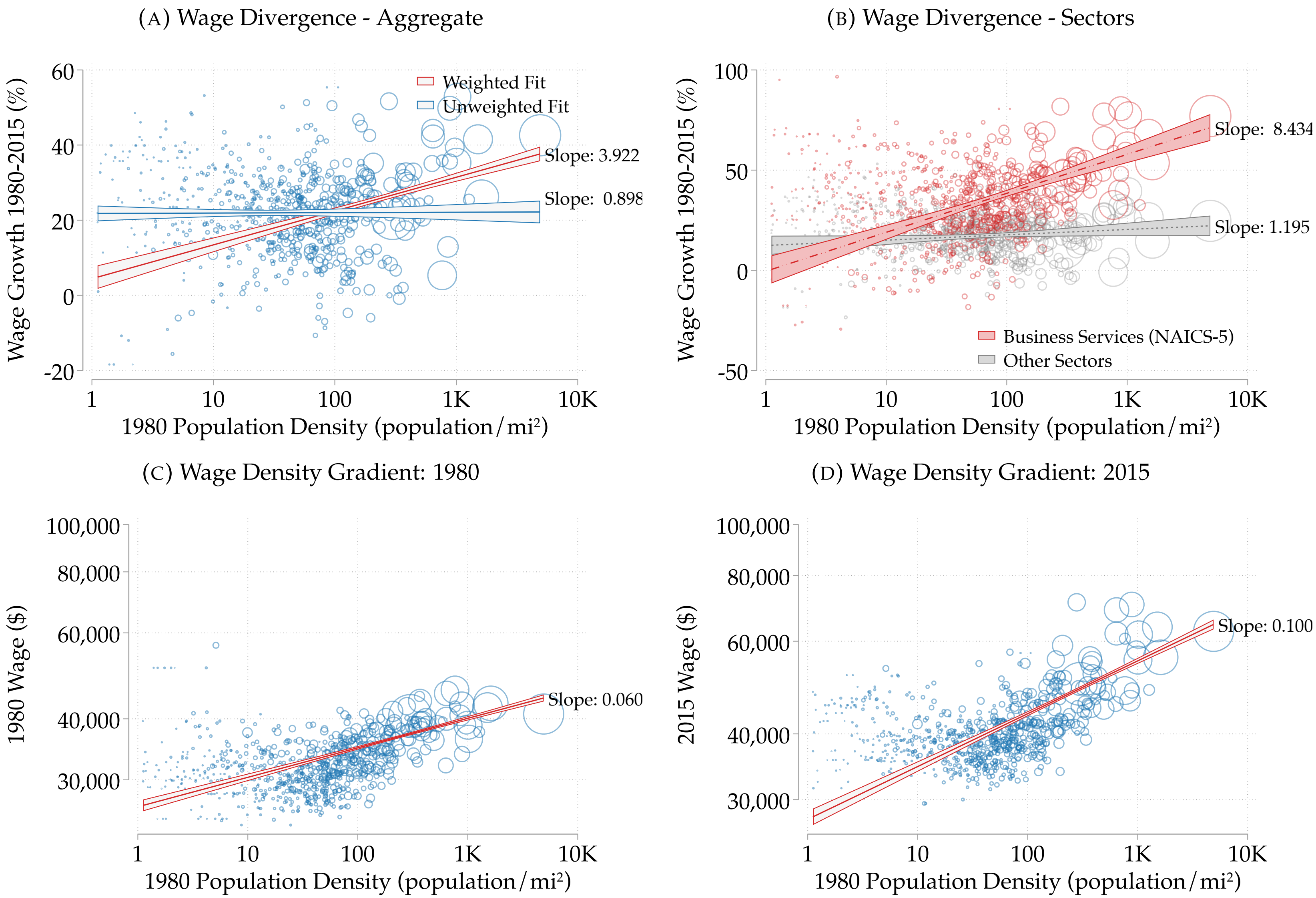


FIGURE SM.3: WAGE DIVERGENCE, 1980-2015



Notes: All wages are adjusted by CPI-U to 2015 dollars. Size of the bubble is proportional to total commuting zone population at the start of each period. The data used is the Decennial Census (1970,1980) and the ACS (2015)

FIGURE SM.4: URBAN WAGE GRADIENT 1975-2015



Notes: Panel (A) shows total average wage growth between 1980 and 2015 at the commuting zone level. Panel (B) decomposes this by sector. Panel (C) and (D) shows the relationship between average wage income and employment and the commuting zone level in 1980 (left) and 2015 (right). The data used is the Decennial Census (1980-2010) and the ACS (2015).The data is adjusted by the BLS CPI-U.Size of the bubble is proportional to total commuting zone employment.

C. Sectoral Wage Gradients Across Data Sources

Data from the QCEW and LBD datasets use administrative records; the first using records collected from unemployment insurance filings and the second using data from tax records. In general, these two data sets closely track each other in terms of aggregate wage and employment growth. However, there are some discrepancies between them on the one side and the self-reported data from the Decennial Census and the ACS on the other side. In Figures SM.5a and SM.5b we plot average wages and total employment over time for 4 types of Business Services (NAICS 51, 52, 54, 55) and all other industries in all three data sets. Figures SM.5c and SM.5d also plot wages and employment growth across the groups of commuting zones used throughout the paper.

In particular, there appears to be a much lower employment count for Business Services and a more muted wage premium in the public Census data. We hypothesize that individuals are reporting their firm's sectoral classification, rather than their establishment's sectoral classification.

We explore this possibility in a case study of Fayetteville (AR) commuting zone, which includes the Bentonville headquarters and support facilities of the largest American retailer, Walmart.[48] In the 2015 QCEW, using administrative records, retail accounts for 12% of employment in the Fayetteville, AR metro area. This is broadly in line with the national average, which stands at 11% in the same dataset. Wages are also broadly in line, with retail workers in Fayetteville making 10% less than the national average for retail workers, and making 45% less than the average worker across industries in Fayetteville.[49] The data for Fayetteville looks very different in the self-reported ACS sample for 2015. In the Fayetteville CZ, retail accounts for 20% of employment, double the national average. Wages in Fayetteville retail are 45% higher than retail workers in the rest of the nation, and 30% higher than all other workers in Fayetteville. Also retail shows a disproportionately high college and post-graduate share of employment, with a lot of workers in legal, accounting, and management roles.

[48] We will only use public-use data from the QCEW here to maintain privacy of tax records from the LBD. The Fayetteville (AR) commuting zone contains both the towns of Fayetteville and Bentonville.

[49] Note, the QCEW suppresses the statistics for "Management of companies and enterprises" in the Fayetteville area as it would effectively reveal the average wage for a single firm or establishment. Public data from the town's chamber of commerce seems to indicate that Walmart must be this firm/establishment.

This provides evidence for our hypothesis that workers are using their firm's industry (retail for Walmart), rather than their establishment (corporate headquarters for Walmart).

We further explore difference between datasets in the supplemental materials.

D. Aggregate and Density Trends Across Data Sources

In this online appendix, we decompose in detail the trends underlying both the wage and sectoral effects in section 2. We explore robustness with respect to data sources and variation in measure of population density and population size.

First, we plot employment shares and wages by sector over time in figures SM.6 and SM.7. In Panel SM.6a, we show how employment in business services has grown in employment from 1980 to 2015, along with employment in the education and medical sectors. We also can track the decrease in the manufacturing sector, decreasing from 23% to 8% employment shares.

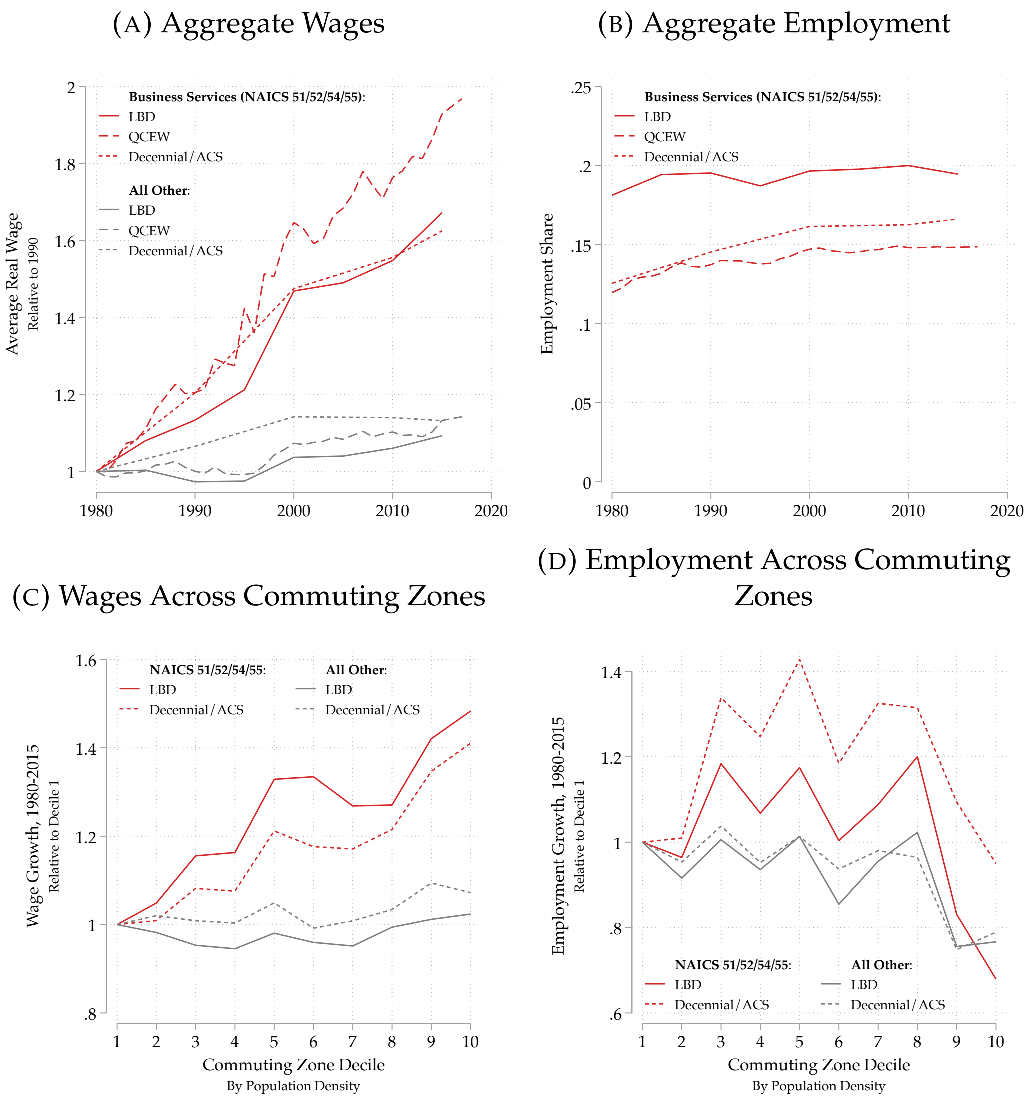
In Panel SM.6b, we track changes in real wages by sector from 1980 to 2015. Business services saw substantial wage growth, outpacing all other sectors. Panel SM.6c shows the relative growth, business service workers saw 70% growth in real wages. No other sector saw more than 35% growth.

Results in Figure SM.7 decompose the geographic aspect of this wage and sectoral growth. Panel SM.7a shows wage growth across space from 1980-2015, normalized to growth in the least dense commuting zones (the bottom 10% by density in 1980). Business services saw highly biased wage growth. Those in dense commuting zones, such as New York, saw 40-50% higher wage growth than those in more rural locations. Other sectors, such as manufacturing and government had modestly urban-biased growth. Other large sectors, such as education, medical, and hospitality saw no urban-biased growth. Natural resource sectors saw rural biased growth.

The above figures are from the Decennial Census, for confirmation, we replicate Figures SM.7a and SM.6c across datasets.

In SM.8 we compare aggregate wage trends in three datasets by broad NAICS categories. The top panels shows QCEW, middle shows the LBD, and the bottom panel uses the Census/ACS. In all three, business services significantly outpaces all other sectors. Both the QCEW and Census/ACS have relatively

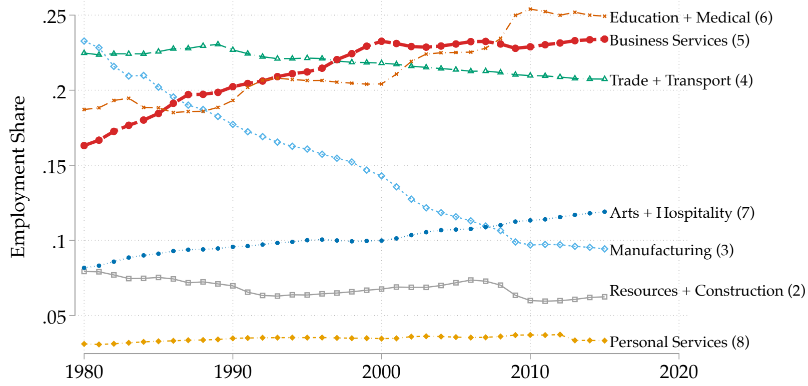
FIGURE SM.5: COMPARING BUSINESS SERVICES WAGE GROWTH ACROSS DATA SETS



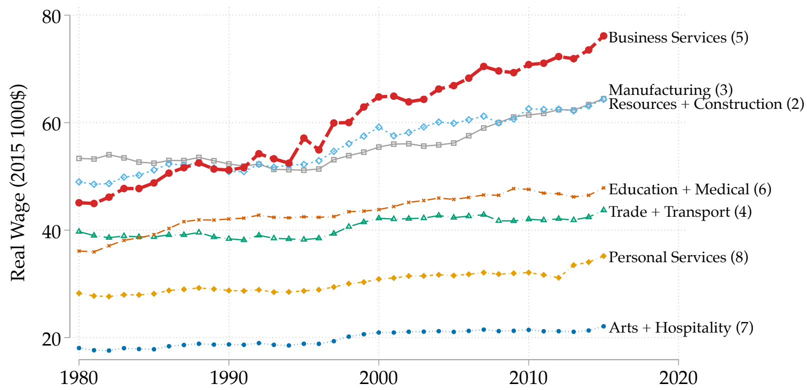
Notes: This is a comparison of four administrative and survey data sources, the Longitudinal Business Database (LBD), the Quarterly Census of Employment and Wages (QCEW) and the combined Decennial Census (Decennial) and American Community Survey (ACS). Panel (A) highlights different wages across Business Services and other sectors over time. Panel (B) highlights the employment share of the Business Services sector over time. Panels (C) and (D) highlights the spatial wage gradient change and employment change for both sectors from 1980-2015 for the LBD and the Decennial/ACS Data. Non-censored QCEW data is unavailable spatially due to disclosure risk.

FIGURE SM.6: AGGREGATE SECTORAL TRENDS

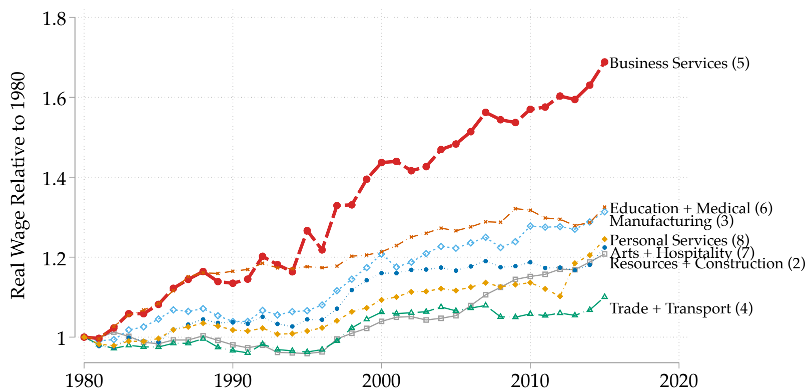
(A) Aggregate Employment Shares



(B) Aggregate Real Wages



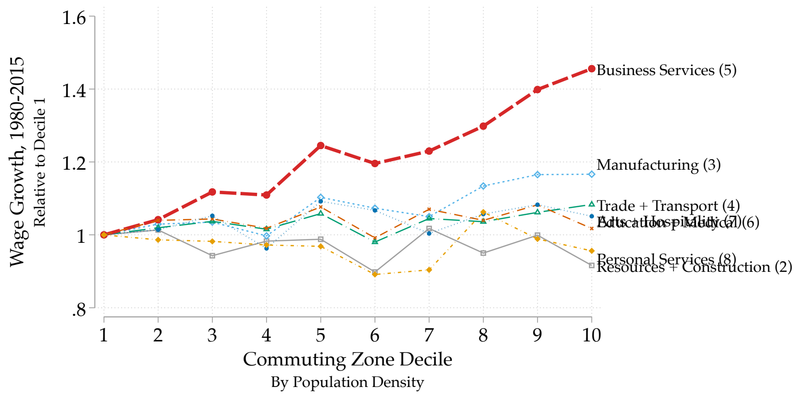
(C) Relative Wages



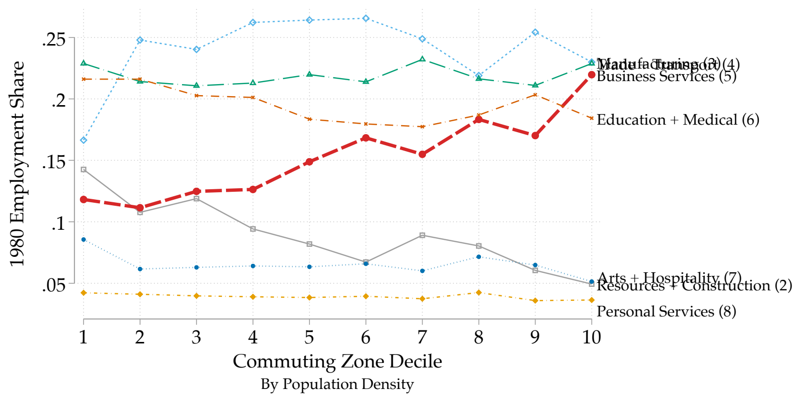
*Notes:*Panel (A) highlights the employment share of the Business Services sector over time. Panel (B) highlights different wages across Business Services and other sectors over time. Panel (C) Normalizes wages to 1980 levels. Source: Quarterly Census of Employment and Wages (QCEW)

FIGURE SM.7: RELATIVE COMMUTING ZONE GROWTH

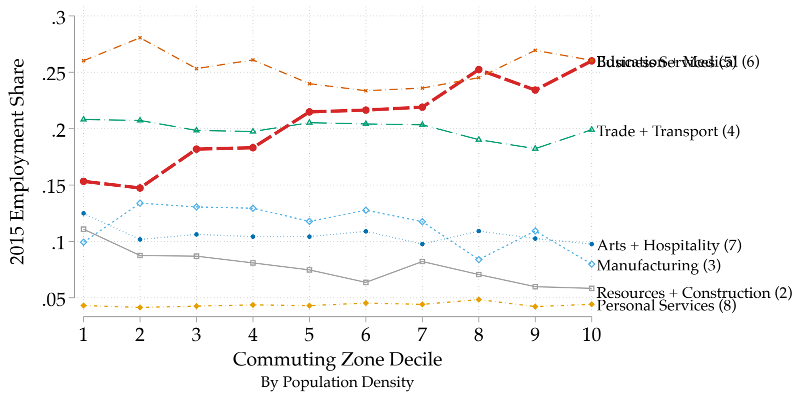
(A) Wage-Density Gradient



(B) Employment Shares - 1980



(C) Employment Shares 2015



Notes: Panel (A) highlights different wages across Business Services and other sectors over time and space. Panel (B) and (C) highlights the employment share of the Business Services sector over time. Source: Decennial Census/American Community Survey full coverage of the economy (though both have issues with self-employment income and agriculture), though data in the Census/ACS may suffer from issues due to the self-reported nature of employment and income. The LBD only covers private sector employment. Overall, across all three datasets, average real income in business services increased 50% or more in our time period. The next closest sector (education and medical services) ranged between 25% and 30%. Some sectors saw sub-10% wage growth over 35 years, with trade, and transport lagging behind other sectors. Due to coverage issues, agriculture and government are excluded.

In SM.9 we compare spatially/urban-biased wage growth across our three datasets. The top panel uses the QCEW. However accurate industry-commuting zone level data is missing before 1990, so we compare wages for the 25 years from 1990 to 2015. This data requires caveats as smaller commuting zones may have missing observations due to privacy concerns. As before, business services is the only industry with a significantly urban-biased wage growth. Manufacturing, arts, and hospitality are all stable across urban and rural commuting zones. All other sectors saw slower wage growth in urban areas than in rural areas, with personal services, trade, and transport exhibiting significant negative urban biases. The heavily skilled education and medical sectors also show pronounced rural-biased wage growth.

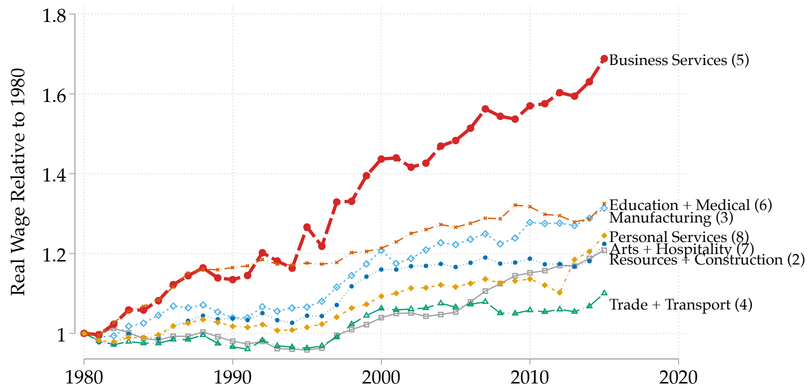
The middle panel replicates this using the LBD. Due to Census disclosure reasons, we present the industry decomposition dividing business services in to "skilled" categories, grouping NAICS 51, 52, 54, and 55 together as "Skilled Scalable Services" and including NAICS 53 and 56 (housing, rental, administrative, support and waste services) with personal services. As before, the large business services/skilled scalable services category exhibits significant urban bias in wage growth. The similarly educated education and medical services sector shows a negative urban bias in wage growth. The bottom panel replicate this using the Census/ACS.

In Figure SM.10 we compute our spatial gradients using three different measure of density/size. The top panel SM.10a simply replicates the figure from SM.9 using population density computed over an entire commuting zone. In SM.10b, we replicate this using the total population of a commuting zone.

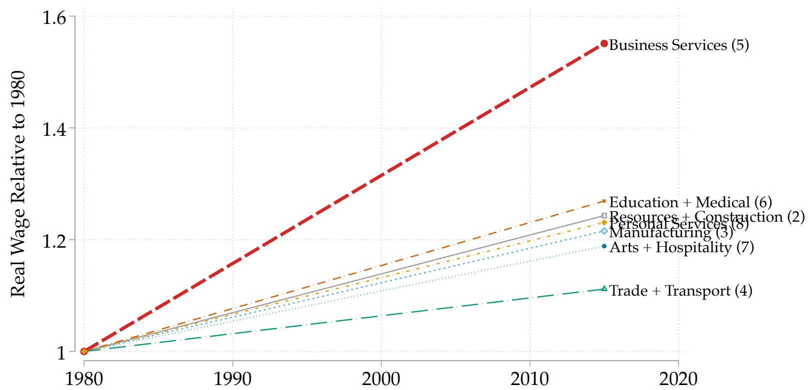
In SM.10c we use the 1980 tract-weighted density of a commuting zone. We construct this density by considering the density of each census tract (a collection of a few census block groups) and creating an aggregate commuting-

FIGURE SM.8: AGGREGATE WAGE GROWTH

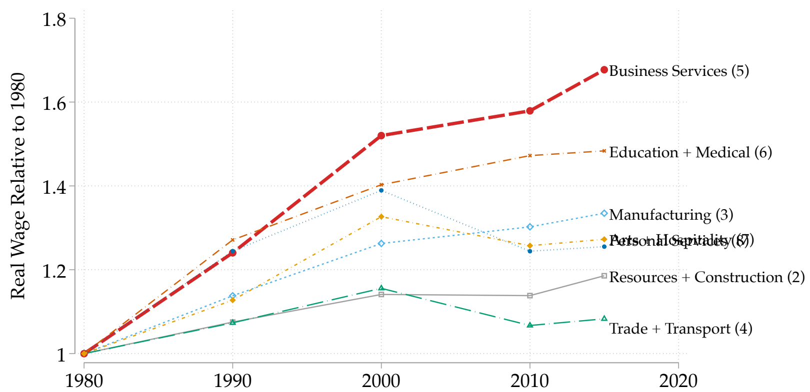
(A) Wage Growth - QCEW



(B) Wage Growth - LBD



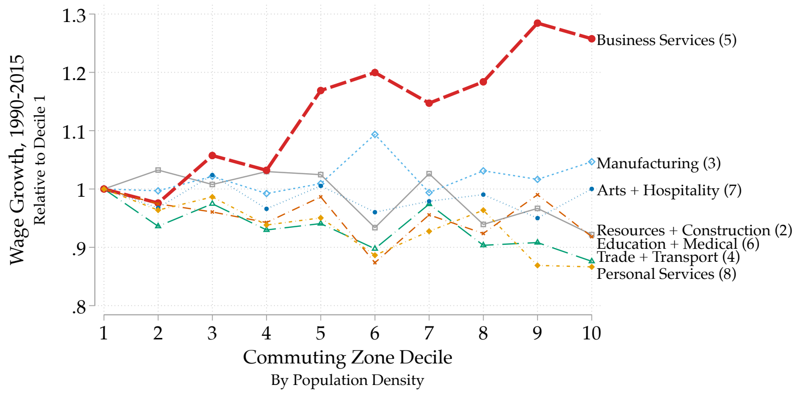
(C) Wage Growth - Census/ACS



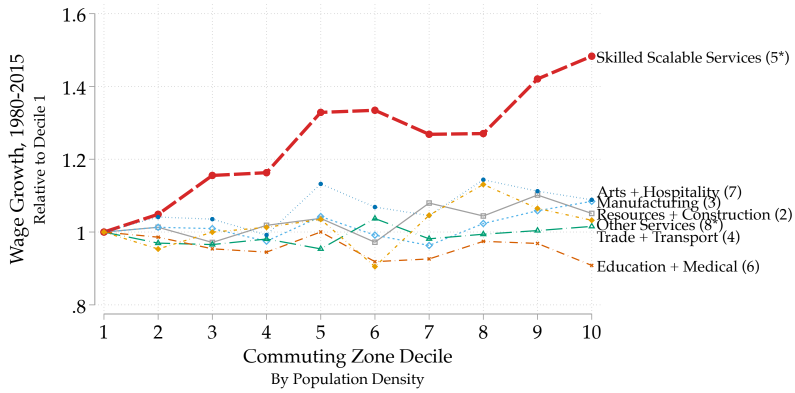
Notes: See text for detailed description.

FIGURE SM.9: WAGE GROWTH GRADIENT

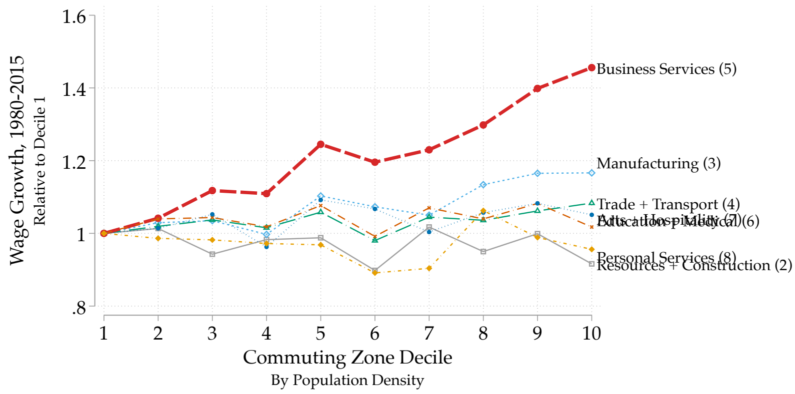
(A) Wage Growth - QCEW



(B) Wage Growth - LBD



(C) Wage Growth - Census/ACS



Notes: See text for detailed description.

zone density by taking the population weighted mean of every tract. This de-emphasizes rural tracts and empty land, like much of the edges of the LA commuting zone. In 1980, some rural areas were not assigned tracts, we simply substitute the county-level density for these remote and rural locations.

D.1 Private versus total Employment

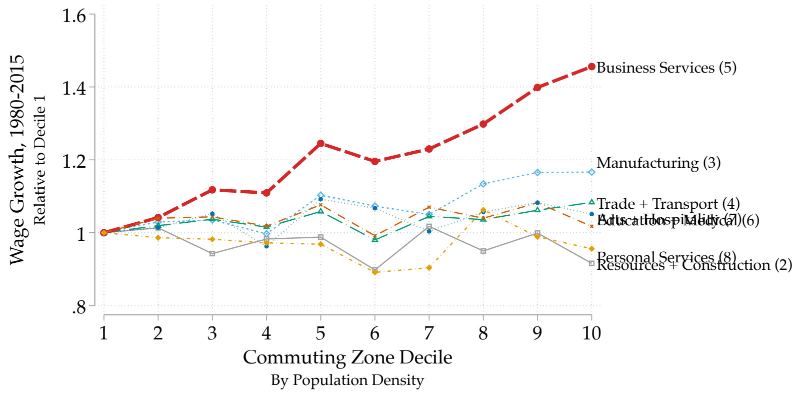
In Figure SM.12 we separate estimate the wage elasticity for all workers, not just those in private employment. The blue circles replicate figure 2 for private employment. The red circles use all employment. Broadly, both lines similar identical trends.

D.2 Wage Dispersion

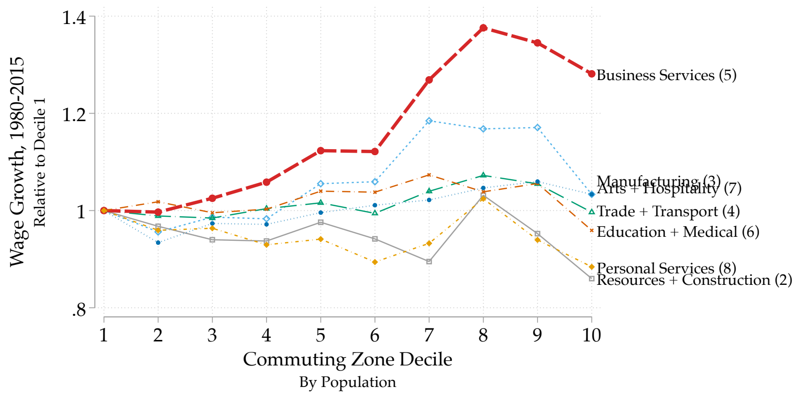
So far, we've simply considered how much wages depend on a commuting zone's size or density. In Figure SM.13, we simply consider the dispersion in the logarithm of commuting zone-level wages across time (weighted by total employment). Using QCEW data for unsuppressed counties, we show a kernel density (with bandwidth .05) for 1980, 1990, 2000, 2010, and 2015. Wage dispersion increases from 1980-2000, with a leveling off from 2000-2015. This mirrors the patterns we observe in the aggregate urban wage gradient.

FIGURE SM.10: WAGE GROWTH GRADIENT - CENSUS/ACS

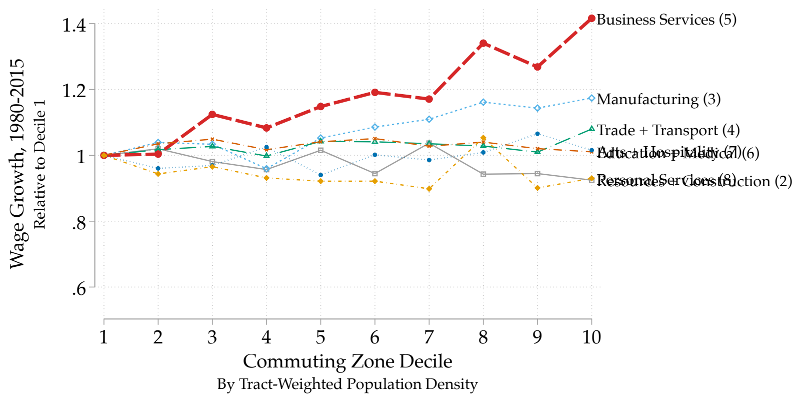
(A) By Density



(B) By Population



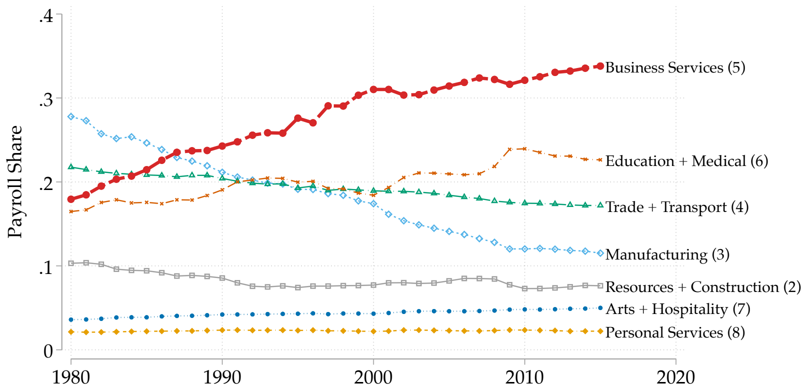
(C) By Tract-Weighted Density



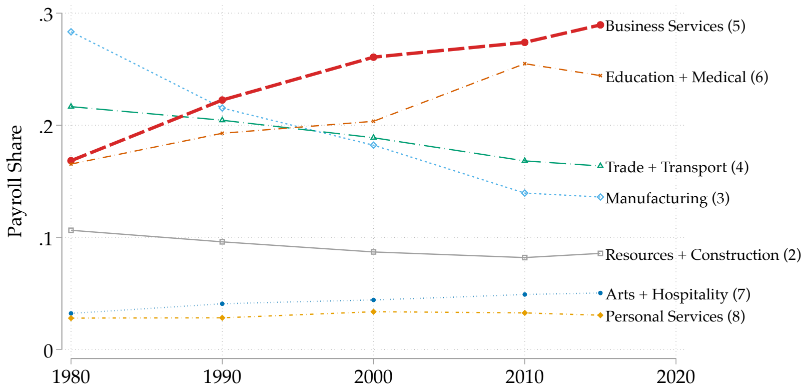
Notes: See text for detailed description.

FIGURE SM.11: PAYROLL SHARES

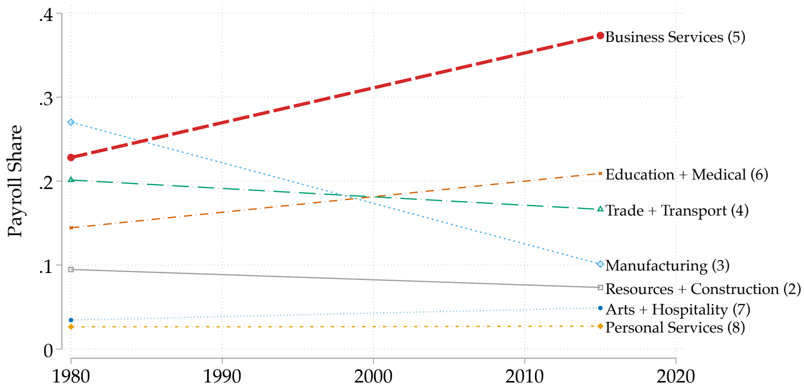
(A) QCEW



(B) ACS

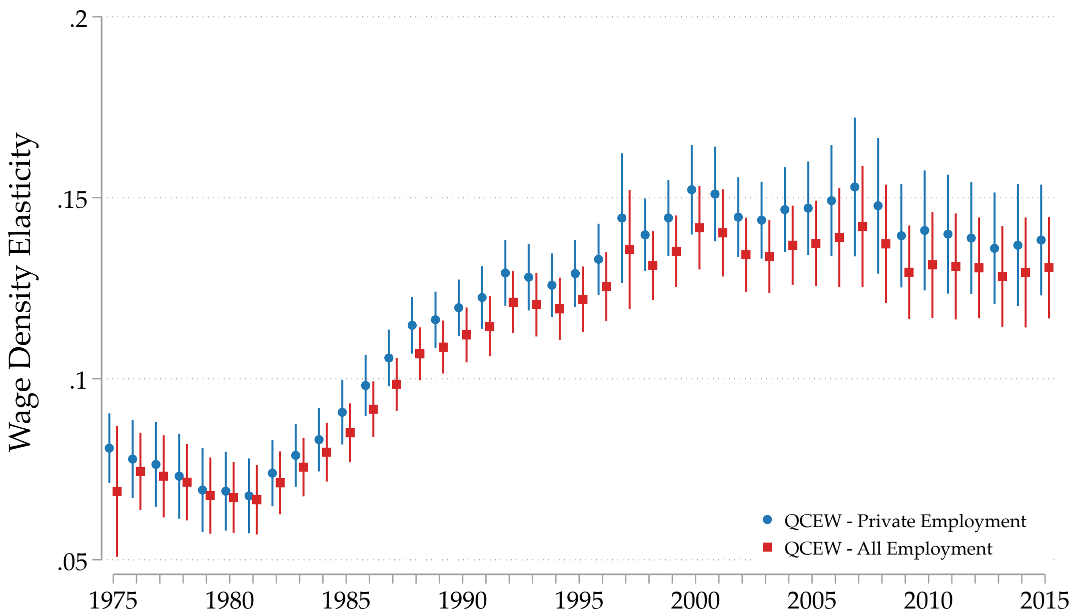


(C) LBD



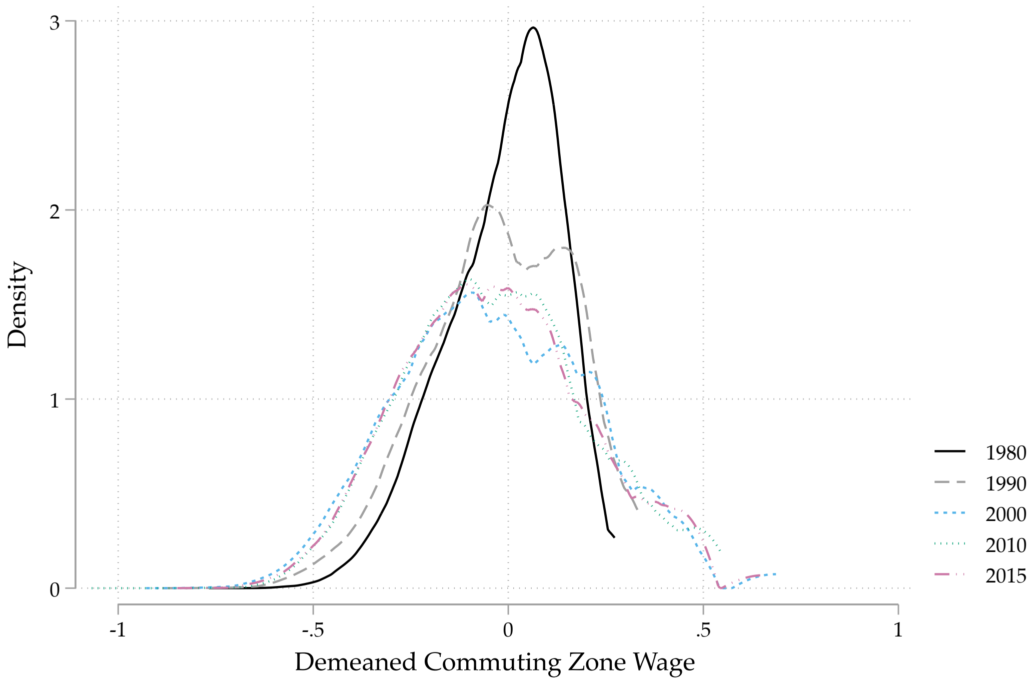
Notes: See text for detailed description.

FIGURE SM.12: URBAN WAGE PREMIUM 1975-2015 - PRIVATE VERSUS ALL EMPLOYMENT



Notes: This figure show the wage-density gradient coefficients β_t across commuting zone, r, for each year from the regression $\ln w_{rt} = \alpha + \phi_t + \beta_t \ln density_r + \epsilon_{rt}$ and the wage-population elasticity substitution $\ln population_r$.

FIGURE SM.13: DISPERSION IN WAGES



Notes: This figure displays the commuting zone (Tolbert and Sizer (1996)) dispersion in the logarithm of wages and income from the QCEW for 1980, 1990, 2000, 2010, and 2015. We demean wages by the national average and use a Epanechnikov kernel function with a bandwidth of 0.05.

D.3 Comparison With IRS Data

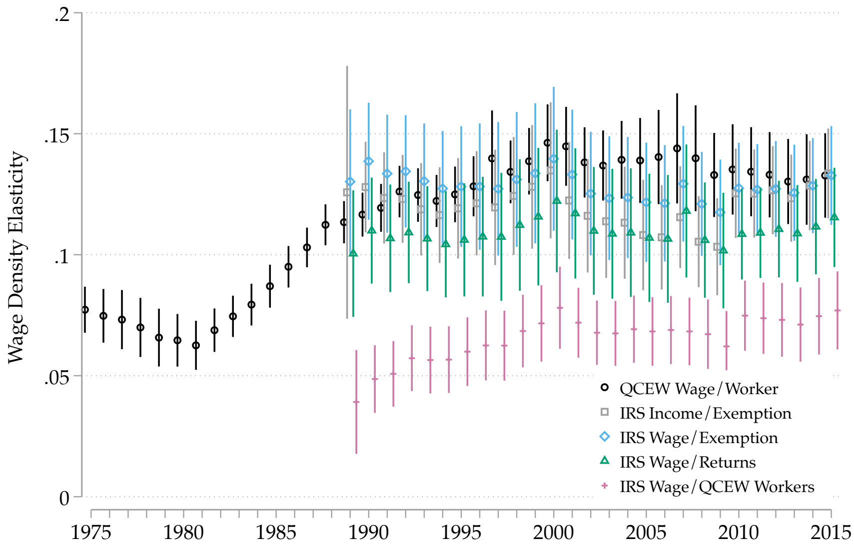
The US Internal Revenue Service (IRS) also provides county data from 1989 to today. Derived from the Statistics of Income (SOI) databases, the IRS data contains the entire universe of tax filers, not just workers. We re-aggregate this data back to the commuting zone level. The IRS also reports non-wage income and wage income separately. In SM.14 we explore four comparisons. The QCEW trendline appears in black circles. The first comparison is with IRS total income (including non-wage income, such as dividends and capital gains) divided by the number of exceptions, which may include the number of dependents. The second comparison consider just considers IRS salary and wage income (after pre-tax deductions such as health care and retirement) divided by the number of exceptions. The third comparison considers wage income, but divided by the number of tax returns (which may be joint between married spouses or individual). [50] Fourth and in the purple crosses, instead of using IRS reported data for the number of workers, we substitute the number of full-time employees from the QCEW.

As shown, the reported coefficients, with robust standard errors, are extremely noisy. The last trend, using the IRS wage data divided by the of QCEW full-time employees, most closely follows the headline QCEW trend, albeit at a lower level.

There are many reasons for the divergence of the first three IRS trends. The first is in the fact that family composition has been undergoing vast changes, which smaller families and fewer married couples; thus breaking any inter-temporal link between the number of tax filings and the number of workers. The last trend does not suffer these issues, as it uses administrative data on the number of workers as the denominator. However it does have a lower level. There are a few potential reasons. First, the QCEW uses data on wages paid by the employers subject to unemployment insurance, which may be different from wages subject to federal income taxation, which has a significant number of allowed deductions. Second, even though we use commuting zones, worker homes and employer locations may differ. Third, some retirement income is classified as salary/wage income and is still subject to standard federal tax. Taken all together, we'd expect and observe an attenuation in the wage-density gradient.

[50] The IRS only reports the number of returns with wage income after 2010. Similar reporting requirements hold for the number of single and jointly filed returns

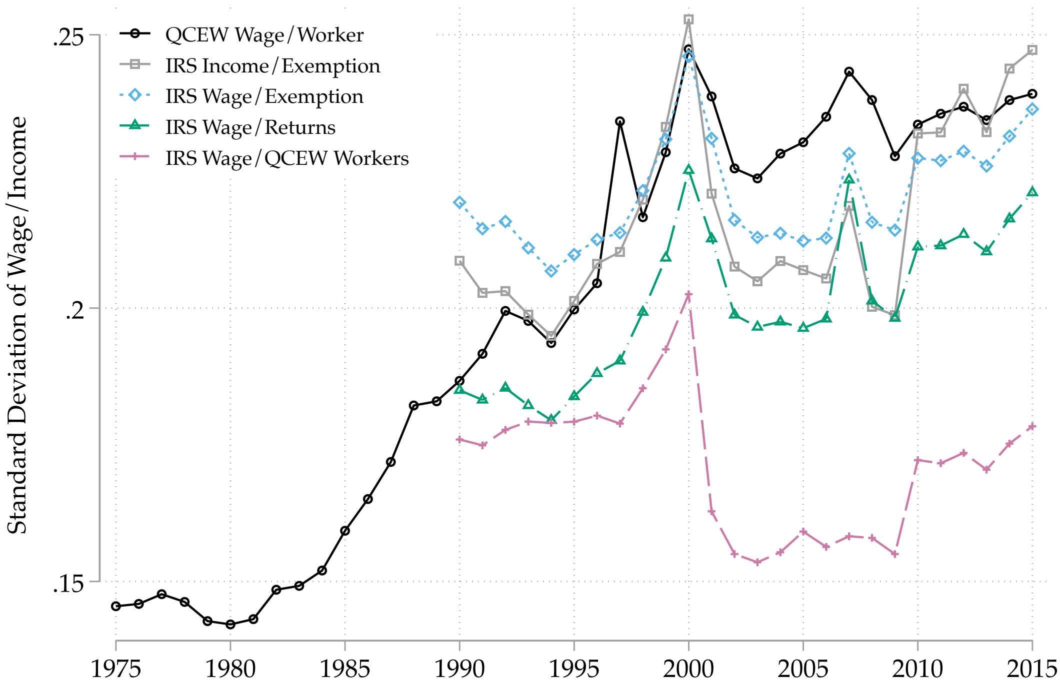
FIGURE SM.14: URBAN WAGE PREMIUM - IRS COMPARISON



Notes: This figure displays the commuting zone (Tolbert and Sizer (1996)) urban wage premium defined 5 ways. This figure show the wage-density gradient coefficients β_t across commuting zone, r, for each year from the regression $\ln w_{rt} = \alpha + \phi_t + \beta_t \ln density_r + \epsilon_{rt}$. The first (circle) uses total QCEW wages divided by employment. The second (square) uses IRS total income divided by the number of exemptions. The third (diamond) uses IRS wages and salary income divided by the number of exemptions. The fourth (diamond) uses IRS wages and salary income divided by the total number of tax returns. The last line (cross) uses IRS wages and salary income divided by the total number of workers from the QCEW.

In figure SM.15 we similarly plot the standard deviation of the commuting zone figure (weighted by total number of QCEW employment) from figure SM.14. The same characteristics and issues from Figure SM.14 are evident.

FIGURE SM.15: STANDARD DEVIATION IN WAGES AND INCOME



Notes: This figure displays the commuting zone (Tolbert and Sizer (1996)) dispersion in the logarithm of wages and income defined 5 ways. The first (circle) uses total QCEW wages divided by employment. The second (square) uses IRS total income divided by the number of exemptions. The third (diamond) uses IRS wages and salary income divided by the number of exemptions. The fourth (diamond) uses IRS wages and salary income divided by the total number of tax returns. The last line (cross) uses IRS wages and salary income divided by the total number of workers from the QCEW.