The Urban Wage Premium in Historical Perspective*

Kyle Butts† Taylor Jaworski‡ Carl T. Kitchens§

December 5, 2024

Abstract

We estimate the urban wage premium in the United States from 1940 to 2010. Drawing on recent advances in the literature on selection on unobservables, we show how to control for heterogeneity in the characteristics of individuals that choose to live in cities to address endogenous sorting. Estimates from naive comparisons of individuals living in urban versus rural areas substantially overstate the urban wage premium. We find that the premium is highest in the middle of the twentieth century (about 12 percent in 1940 and 1950) relative to the early in twenty-first century (declining to a few percent by mid-2020). Overall, the urban wage premium is decreasing and sorting explains a larger fraction of the difference in urban versus rural earnings across our sample period.

*This version of the paper includes preliminary results. Please do not cite without permission.

†University of Arkansas, Department of Economics, (kyle.butts@colorado.edu)

‡CU-Boulder, Department of Economics, and NBER (taylor.jaworski@colorado.edu)

§Florida State, Department of Economics, and NBER (ckitchens@fsu.edu)

1 Introduction

The average U.S. worker living in an urban area in 2010 earns 35 percent more than their non-urban counterparts.[1] This pattern of higher average wages in cities holds across the globe and through out the past 80 years (Boustan et al., 2013). The fact that wages are higher in cities cannot be interpreted as the causal effect of cities due to the potential for non-random sorting of people or firms into cities based on preferences, skills, or productivity.[2] Is the difference driven by increased productivity stemming from agglomeration economies or high-ability workers sorting into urban areas (e.g. Behrens, Duranton and Robert-Nicoud, 2014; Moretti, 2011; Puga, 2010)? Empirical evidence focused in the last 30 years has found a large portion of the observed wage premium is driven by non-random sorting, but a true effect of urbanity on wages still exists.

This paper asks how has selection and the true causal wage-benefits of urban living has changed over time. Past research shows that the observed raw differences in wages between urban and non-urban workers was around 35 percent in 1940, decreasing by one-third to one half in 1980 to 20-25 percent, and increasing again to 35 percent by 2010 (Dauth, Findeisen, Moretti and Suedekum, 2022; Boustan, Bunten and Hearey, 2013). However, no estimates exist prior to 1990 on whether this is due to changes in sorting patterns or changes to the causal returns to urbanity. Given the role of agglomeration in economic models of cities, regions, and national growth, it is important to understand long run changes in the causal effect of cities on wages. Moreover, understanding how the underlying drivers of selection and strength of agglomeration forces have varied over time is important for predicting the success of place-based policies (e.g. Kerr and Robert-Nicoud, 2020) or to understand growth trajectories in the context of endogenous growth models (e.g. Martin and Ottaviano, 2001).

Research in this area has typically addressed these concerns using detailed worker- and firm-level panel data. Comparing the same individuals before and after moving to an urban

[1] We use the 2010 ACS estimates to compute the average annual wage for individuals living in MSAs and outside of MSAs based on individuals that are in the labor force, report non-negative wage income, and who are not missing wage income data.

[2] For example, positive selection on the firm side may come from more productive firms paying higher wages, and only high-productivity firms locating and surviving in cities where larger markets induce tougher competition.

area removes time-invariant wage variation due to worker characteristics, such as ability. Then, the remaining gap or subsequent wage growth between urban and rural earnings is attributed to agglomeration forces stemming from knowledge sharing, improved firm-worker matching, human capital accumulation, labor market pooling, or input-output linkages (Baum-Snow and Pavan, 2012; Combes, Duranton and Gobillon, 2008; Dauth, Findeisen, Moretti and Suedekum, 2022; Glaeser and Maré, 2001; Gould, 2007; Yankow, 2006). The previous literature provides a range of estimates regarding the urban wage premium. For example, Glaeser and Maré (2001) estimate an urban wage premium between 4.5 and 11 percent after controlling for individual characteristics and, more importantly, unobserved heterogeneity with worker-specific fixed effects. Gould (2007) finds a premium of 11.5 percent for white-collar workers and a smaller premium of 1.2 percent for blue-collar workers.[3]

Due to limited data availability the existing literature has focused primarily on the last three decades and little is known about the evolution of the urban wage premium before 1980. The primary challenge is that census data only provides a cross-sectional snapshot of who lives in an urban area and how much they earn, preventing the standard strategy of tracking workers when they move into or out of an urban area. Hence, any cross-sectional estimates comparing urban to non-urban workers will be confounded by unobserved worker characteristics that drive sorting.

In this paper, we show how to use readily available census data to estimate the urban wage premium while controlling for observed individual characteristics as well as unobserved characteristics following the approach in Altonji and Mansfield (2018). Their setup is based on the classic urban location choice model. Workers vary in their willingness to pay for location-specific amenities based on observable and, importantly, *unobservable* characteristics. Because of differences in willingness-to-pay, location-specific averages of observable characteristics vary based on a location's underlying amenities. Similarly, the averages of unobservable characteristics vary because of (potentially different) location-specific amenities. If the amenities driving both observable and unobservable sorting are the same, the model

[3] Rosenthal and Strange (2004) and, more recently, Diamond and Gaubert (2021) survey the large literature that describes the theoretical underpinnings and empirical findings related to agglomeration and urban sorting.

suggests that controlling for observable location-specific averages is sufficient to control for unobservable averages.[4] This approach is crucial as it allows us to address sorting when estimating the historical urban wage premium without individual-level longitudinal data.

We estimate the urban wage premium from 1940 to 2010.[5] Focusing first on the period since 1980, where we can make direct comparisons to the existing literature, we document an urban wage premium starting at 7.7 percent in 1980 and declining to around 5 percent in the following decades. This is consistent with the results reported in Glaeser and Maré (2001), Yankow (2006), and Baum-Snow and Pavan (2012) and provides support that our approach (drawing on Altonji and Mansfield (2018)) addresses the key identification concern related to sorting. Second, moving to earlier decades, we estimate an urban wage premium of 12.9 percent in 1940, 16.2 percent in 1950, 13.3 percent in 1960, and 9.2 percent in 1970. The pattern suggests that agglomeration forces were relatively more important prior to 1980, although sorting still drives a large portion of the observed premium between 1940 and 1980.[6]

We find that the causal urban wage premium, which reflects the fact that workers moving to cities earn more, was large in 1940 but steadily declined over time, reaching its lowest point by 2010, the most recent year in our data. This pattern aligns with several well-documented trends in the literature. First, the increasing role of sorting in explaining the wage differential between urban and non-urban areas is consistent with the rising migration of highly educated workers to cities (Shapiro, 2006; Moretti, 2013; Diamond, 2016). Though raw wage premiums have been growing in recent years, this migration and our results suggest this is a compositional effect rather than a growth in the causal premium. This trend is also linked to structural changes in manufacturing after 1950 (Dumais, Ellison and Glaeser, 2002; Glaeser and Ponzetto, 2010), which contributed to shifting economic activity toward urban areas. Historically, cities were manufacturing hubs with large agglomeration forces which enhanced worker's productivity. The shift away from manufacturing in cities could be a

[4] We discuss this assumption below and provide evidence that builds confidence in the plausibility.

[5] Specifically, we use only repeated cross-sections from the US decennial censuses and American Community Survey beginning in 1940 when information on individual wages was first reported.

[6] Due to changes in the definition of cities over time (smaller cities being added over-time), we also estimate the urban wage premium on consistent set of locations based on the urban status in 1940. The results remain very similar.

driver of the decline in the causal wage premium. Moreover, the increased mobility of highly skilled workers in response to tax policy further supports these findings (Moretti and Wilson, 2017; Verginer and Riccaboni, 2021; Akcigit, Baslandze and Stantcheva, 2016).

Another contributing factor is the growing body of evidence that highly skilled workers are increasingly sorting into cities based on improved local amenities (Baum-Snow and Hartley, 2020; Diamond, 2016; Su, 2022; Glaeser, Kolko and Saiz, 2001; Glaeser and Shapiro, 2003; Glaeser and Gottlieb, 2009). [7]

The first concern with our results is that while urban workers might earn more in wages, they also face larger housing costs and hence there might not be a *real* benefit to urbanity after adjusting for housing costs. We calculate a wage meausre that subtracts off housing costs following Ganong and Shoag (2017), and reestimate both the raw and causal (net-)wage premium. After adjusting for housing costs, the raw wage premium is decreased significantly between around 18% in recent years and . The decline is smallest around 1970 with around a 5 percentage point decline and larger at the beginning and end of our sample with around 15 percentage point decline. However, our causal estimates remain roughly unchanged. There is a 3-4 percentage point decline in 1940 and 1950, but otherwise our estimates remain stable.

In addition to the main findings, we explore several potential sources of heterogeneity. First, we examine changes in the urban wage premium across the distribution of city sizes. To do so, we estimate wage premiums separately for the 20 largest urban areas (based on 1940 populations) and other urban areas. We show that in 1940, the largest cities had larger urban wage premiums relative to smaller urban areas. However, the gap has narrowed since the 1980s. The pattern we document is consistent with the transition of cities from heavy manufacturing bases to idea-generating locations.

Second, we consider regional differences and confirm that urban wage premium is falling *within* each region. The consistent fall of the premium within each region suggests that it

[7] Our results are also consistent with a growing literature that emphasizes the potential to overstate agglomeration effects in small samples (Ahlfeldt, Albers and Behrens, 2022; Dingel and Tintelnot, 2020; Schoefer and Ziv, 2022). For instance, Schoefer and Ziv (2022) show that as much as three-quarters of the variation in regional productivity dispersion may be due to small sample bias. Likewise, Dingel and Tintelnot (2020) illustrate that when accounting for granularity, the potential impacts of Amazon's HQ2 are smaller compared to predictions from traditional spatial equilibrium models.

is not changes in regional development (i.e., the deindustrialization of the North and rise of the Sunbelt) driving changes in the urban wage premium. We do find important regional trends that are consistent with the historical narrative. The Midwest had the largest urban wage premia in the middle of the century (around 25% in 1950) that could reflect the strong agglomeration forces of manufacturing employment. The west has had a growth in estimated urban wage premium since 1990, perhaps picking up on the growth of the technology sector in Silicon Valley and other regions.

Third, we estimate urban wage premium for college and non-college educated workers. We find that throughout our sample, the urban wage premium is higher for college-educated workers. This is consistent with the complentary between skilled workers and cities emphasized by Moretti (2013).

Fourth, we estimate urban wage premium separately for Black and White workers. In all years, the premium for Black workers is no smaller than for White workers and is particularly larger between 1950 and 1970. The Black urban wage premium is estimated to be five to ten percentage points larger in the middle of the 20th century which is in line with the timing of the Great Migration (Derenoncourt, 2022).

Fifth, we try to understand the role of urban *wage-growth* premium as first documented by Baum-Snow and Pavan (2012). We are fundamentally limited by cross-sectional data and can not observe directly whether a worker's wage grows faster in an urban area. However, we estimate separate wage premia for older and younger workers assuming that the former has more years of urban experience and should presumably have larger wage premia. We indeed find that older workers have higher premia in most years, but the gap shrinks to zero in later years.

Finally, we provide evidence that the decline in the urban wage premium is consistent with the pattern of national wage compression documented by Goldin and Margo (1992).

2 Data and Summary Statistics

In this section, we describe the data we use to estimate the urban wage premium and key empirical patterns regarding the changes in urbanization, education, and the urban versus non-urban wage difference in the second half of the twentieth century. First, we draw on publicly available data covering the period between 1940 and 2010 (Ruggles et al., 2022a,b). For 1940, we use the complete count of the US Census to measure the average earnings per worker for wage earners. For 1960 and 1970, we use the 5 percent samples of the census. For 1950, 1980, 1990, and 2000, we draw on the 1 percent samples of the census. Finally, for 2010, we use the American Community Survey. In each year, we restrict the data to include adult males living in the continental US between the ages of 25 and 65 who report being employed for at least 26 weeks in the sample year. We use person weights to calculate summary statistics and estimate regressions.

To construct average weekly earnings, we divide the reported total income by the number of weeks worked in the survey year.[8] We then convert all earnings to 2015 dollars using the Consumer Price Index. Since housing costs are higher in urban areas, we construct a measure of 'real wages' following Ganong and Shoag (2017). To do so, we use either (i) five percent of the reported housing value or twelve times the reported monthly rent and divide by 48 to get the weekly housing costs. We winsorize the estimated housing cost at the first and 99th percentiles and subtact this from the weekly wages.[9]

In addition to wages, across all samples, consistent information is available on whether the individual is living in an urban area (metropolitan statistical area), age, educational attainment, and race. We use this information to construct the individual- and group-level controls for our empirical analysis. We assign workers to a city based on the metropolitan statistical area (MSA) in which they reside; workers not living in a city are assigned to their state of residence.

We begin by describing a few key empirical patterns in Table 1. First, we report the raw observed difference between urban and rural wages in column 2. Here we document a clear U-shape in the summary statistics. The urban wage gap started at 31.3 percent in

[8] For individuals earning the top income code, we set their total income to the lower top income cutoff multiplied by 1.5. These cutoffs are reported in the IPUMS documentation for the `INCWAGE` variable. In 2015 dollars, the cutoffs are \$84,500 for 1940, \$99,2000 for 1950, \$199,500 for 1960, \$309,000 for 1970, and \$225,000 for 1980. Afterward, IPUMS reports income as the median wage above that year's cutoff. Results are robust to dropping all individuals earning the top income code.

[9] Since 1950 does not have microdata on rent, we use the average rent from the Census tables.

Table 1: Summary Statistics

			% With College Degree		
Year	% Wage Gap	% Urban	Overall	in Urban Areas	in Non-urban Areas
(1)	(2)	(3)	(4)	(5)	(6)
1940	31.3	66.0	8.4	8.8	7.6
1950	21.9	68.7	8.0	8.6	6.7
1960	23.6	65.6	11.6	12.9	9.1
1970	23.6	74.7	16.6	17.8	13.0
1980	19.9	72.8	24.2	26.6	17.5
1990	25.6	73.5	28.0	31.1	19.6
2000	31.4	78.7	30.5	33.4	19.8
2010	30.1	79.4	33.5	36.4	22.2

Notes. This table presents summary statistics for the final ACS sample. The data cleaning process is described in Section 2.

1940 and declined sharply through 1980. Beginning in 1980, the urban wage gap increased until 2000 where it reached nearly the same level as reported in 1940. Next, in column 3, we report the share of the population living in urban locations each year. In 1940, 66 percent of individuals lived in urban areas. By 2010, the urban share increased to 79.4 percent. The pattern of increasing urbanization is consistent with prior literature such as Moretti (2013) and Boustan et al. (2013). To highlight the potential role of sorting, the remaining columns provide information the overall fraction of the population with college degree in each year and differences in this share across urban and non-urban areas. From column 4, the share of the population with a college degree increased over time. From columns 5 and 6, the share of the population with a college degree increases in both urban and rural areas over time; however, the share rises faster in cities. Together, the trends in Table 1 highlight that since 1940 cities have become more populated and more educated relative to rural areas, which suggests increased sorting based on skills. The goal of the empirical approach is to address the potential role of changes in the extent of sorting based on education, other demographic characteristics, and unobserved productivity over time.

The geography of high-wage urban centers has also changed over time. Table 2 reports the

Table 2: Top Ten Cities by Urban Wage Premium in 1940 and 2010

1940		2010	
Name	Premium	Name	Premium
Flint, MI	47.9%	San Jose-Sunnyvale-Santa Clara, CA	63.2%
Detroit, MI	46.0%	Bridgeport-Stamford-Norwalk, CT	54.7%
Washington DC, MD/VA/WV	45.7%	San Francisco-San Mateo-Redwood City,CA	52%
San Francisco, CA	44.0%	Washington-Arlington-Alexandria DC-VA	42.9%
Seattle-Everett, WA	41.7%	Boston-Quincy, MA	42.7%
Lansing-East Lansing, MI	41.0%	Seattle-Bellevue-Everett, WA	42.2%
New York, NY	40.2%	Santa Cruz-Watsonville, CA	36.7%
Sacramento, CA	39.2%	Trenton-Ewing, NJ	36.1%
Chicago, IL	38.8%	Midland, TX	34.0%
Rochester, NY	38.1%	Baltimore-Towson, MD	33.6%

Notes. This table presents the top ten cities in terms of observed wage premia relative to non-urban areas.

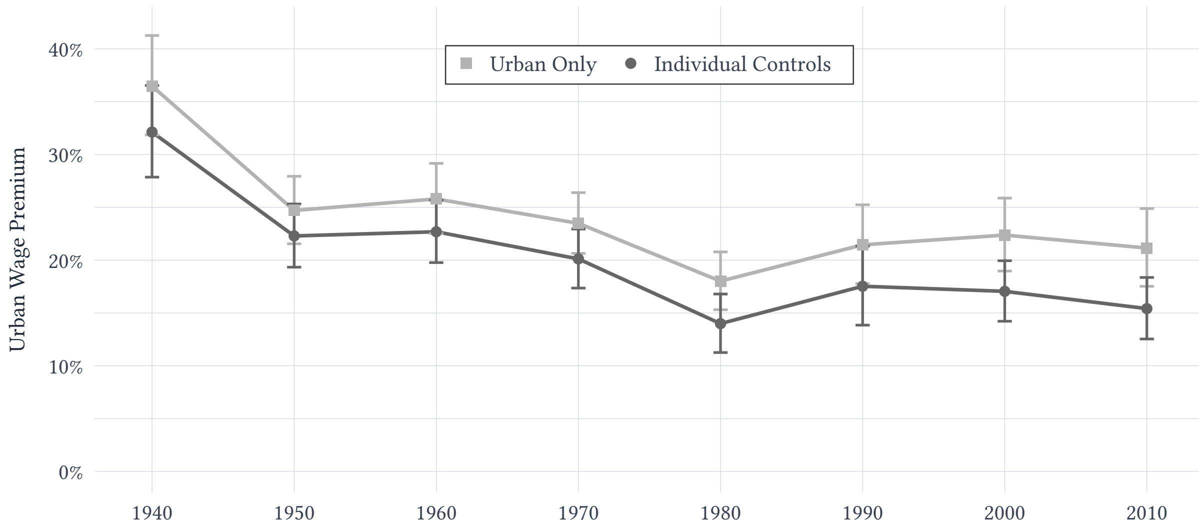
ten urban areas with the highest urban wage differences in 1940 and 2010. In particular, we calculate the difference in the average log wage in an urban and non-urban area. While several of the cities are among the highest paying in both 1940 and in 2010 (i.e., San Francisco, Washington, D.C., San Jose, Seattle), centers of manufacturing in the industrial North and Midwest (i.e., Flint, MI, Detroit, MI, Lansing, MI, Chicago, IL, Rochester, NY) have lost their prominence in favor of cities closely associated with high-tech manufacturing. This descriptive pattern reflects the combination of several changes in the US economy since 1940. The main goal of remainder of the paper is to quantify and better understand changes in the urban wage premium over time and the relative importance of agglomeration (i.e., the causal effect of cities) versus sorting.

3 Measuring the Urban Wage Premium

We begin by estimating the urban-rural wage difference in a simple regression framework. We follow the literature by adding individual-level covariates to address sorting into urban locations. In particular, we estimate the following equation separately for each year:

$$\log(\text{wage}_i) = \beta_0 + \beta_1 \text{urban}_i + \boldsymbol{X}_i \delta + \varepsilon_i, \tag{1}$$

Figure 1: Urban Wage Premium Controlling for Individual Characteristics, 1940-2010



Notes: This figure plots the coefficient for survey-year level estimates of the urban wage premium from equation (1) with different level of controls. The light gray line with square markers recreates the unadjusted point estimates. The dark gray line with circle markers includes controls for age in 5-year bins, education levels, and an indicator for being White. The points for each line are plotted with the 95 percent confidence intervals.

where wage_i is average weekly earnings and $\boldsymbol{X}_i$ is a set of control variables consisting of indicator variables for five-year age bins, educational attainment dummy variables, and an indicator for race (white versus non-white).[10] Our estimate of β_1 is the average premium in percentage terms that urban workers earn relative to non-urban workers (in figures we use percent by calculating $e^{\hat{\beta}_1} - 1$). Without controls, the coefficient reflects the mean difference in earnings between individuals living in urban and non-urban locations. If the set of control variables fully accounts for sorting, we could interpret β_1 as the causal effect of living in an urban location on earnings. However, there are several reasons individuals sort into cities that are not captured by their observable characteristics. For example, individuals with certain attributes may sort to cities because they value specific consumptive amenities, such as entertainment or dining. If these sorting patterns are correlated with productivity, β_1 would suffer from omitted variables bias. As a result, the estimated β_1 should be interpreted as a correlation.

[10] We are limited in our control variables beacause we want these to remain constant across census years.

The results of estimating equation (1) with and without controls are displayed in Figure 1. The grey line with square markers reports the mean urban wage difference without any covariates included. The mean urban differential was highest in 1940 at around 35 percent and falls to a little under 25 percent by 1970. Starting around 1980, the urban wage difference stabilizes at around 20 percent, where it remains to the present. The medium grey line with circle markers reports estimates that include individual-level controls in the specification. Adding individual-level covariates reduces the wage differential in all years. In 1940, individual characteristics explained roughly 10 percent of the earnings gap between non-urban and urban locations. By 2010, differences in individual-level observables explained about 25 percent of the gap. Relative to prior literature such as Yankow (2006) and Glaeser and Maré (2001), our set of observables explains less of the wage differential; however, we are more limited in our set of observable individual-level characteristics.

Overall, the pattern of wage differentials is similar once individual-level characteristics are included. We also show that individual characteristics explain a larger share of the urban wage differential over time. This suggests that sorting – despite our limited set of observables – is increasing over time and is consistent with the raw data in Table 1.[11]

4 Sorting and the Urban Wage Premium

The urban wage premium estimates from the previous section are potentially biased if sorting causes urban workers and non-urban workers to differ in their underlying productivities. Previous work has leveraged panel data on individuals to isolate the effect of living in an urban area by observing individuals who transition across urban and non-urban locations.[12] Given the cross-sectional nature of our data, we cannot adopt these panel data methods. To address the empirical challenge of worker selection into urban locations in cross-sectional data, we draw on the properties of the canonical location choice model in urban economics.

[11] The reason for the difference in estimates relative to Boustan et al. (2013) who find a U-shaped pattern comes from our use of a regression framework. The method in Boustan et al. (2013) estimates $\log(\hat{\mathbb{E}}[\text{Wage} \mid \text{Urban} = 1]) - \log(\hat{\mathbb{E}}[\text{Wage} \mid \text{Urban} = 0])$, which differs from the regression estimate, $\hat{\beta}_1 = \hat{\mathbb{E}}[\log(\text{Wage}) \mid \text{Urban} = 1] - \hat{\mathbb{E}}[\log(\text{Wage}) \mid \text{Urban} = 0]$. The latter is a more precise estimate for average percent difference in wages.

[12] For a recent example that uses matched employee-employer data, see Frings and Kamb (2022). Additionally, Eckert et al. (2022) utilizes the placement of refugees in Denmark to estimate the returns to urban experience.

As shown by Altonji and Mansfield (2018), this model yields statistics from observable data that we use to isolate the causal wage premium. While much of this exposition follows their work, we present it here to ensure the exposition is self-contained

The location choice model says individuals have an indirect utility function over locations ℓ that depends on a vector of location-specific amenities, $\boldsymbol{A}_\ell$, and the cost to live in that location, P_ℓ:

$$\mathrm{V}_{i\ell} = \boldsymbol{W}_i \boldsymbol{A}_\ell - P_\ell + \varepsilon_{i\ell}. \tag{2}$$

Here $\boldsymbol{W}_i$ is a vector of an individuals' willingness to pay for each amenity.[13]

The willingness to pay parameters are assumed to be an additive function of observable characteristics $\boldsymbol{X}_i$ and unobservable characteristics $\boldsymbol{X}_i^{\mathrm{U}}$ of the worker:

$$\mathrm{V}_{i\ell} = (\boldsymbol{X}_i \boldsymbol{\Theta} + \boldsymbol{X}_i^{\mathrm{U}} \boldsymbol{\Theta}^{\mathrm{U}}) \boldsymbol{A}_\ell - P_\ell + \varepsilon_{i\ell}. \tag{3}$$

Each worker receives a set of idiosyncratic preference draws, $\varepsilon_{i\ell}$ for each location ℓ. As is standard in the literature, we assume the shocks $\varepsilon_{i\ell}$ are mean zero and drawn independently from $\boldsymbol{X}_i, \boldsymbol{X}_i^{\mathrm{U}}, \boldsymbol{A}_\ell$, and P_ℓ. Workers then chose the location that maximizes their indirect utility assuming they take the amenities and prices as given.

In this model, the previous section's urban wage premium estimates will be biased if two things are true. First, in order to have sorting occur, urban locations must offer differ in amenities compared to non-urban locations. Second, the unobservable characteristics that change willingness to pay for these amenities, $\boldsymbol{X}_i^{\mathrm{U}}$, must also affect productivity. For example, Diamond (2016) finds that the movement of high-skilled workers into dense cities is driven in large part by differences in preferences for consumptive amenities. This creates our main identification challenge: if the kinds of amenities that cities are abundant in draw workers with higher productivities, then the estimates of (1) will be biased upwards.

Since individuals are maximizing their indirect utility when choosing where to live, we expect location averages of $\boldsymbol{X}_i$ and $\boldsymbol{X}_i^{\mathrm{U}}$ to both be correlated with (some of) the amenities (in

[13] This multinomial model of utility is common in urban economics (e.g. Kline and Moretti, 2014; Diamond and Gaubert, 2021).

particular, amenities where the corresponding rows of $\boldsymbol{\Theta}/\boldsymbol{\Theta}^U$ are large). That is, there exists some mapping between averages of unobservables $\bar{\boldsymbol{X}}_\ell^{\mathrm{U}} \equiv \mathbb{E}\left[\boldsymbol{X}_i^U \mid \ell(i) = \ell\right]$ and amenities $\boldsymbol{A}_\ell$, say $\bar{\boldsymbol{X}}_\ell^{\mathrm{U}} = f^{\mathrm{U}}(\boldsymbol{A}_\ell)$. The same logic holds for averages of observables $\bar{\boldsymbol{X}}_\ell = f(\boldsymbol{A}_\ell)$.

Under conditions we will discuss below, we can invert the latter equation to solve for amenities $\boldsymbol{A}_\ell = f^{-1}(\bar{\boldsymbol{X}}_\ell)$. Then, plugging this into the the former equation yelds $\bar{\boldsymbol{X}}_\ell^{\mathrm{U}} = f^{\mathrm{U}}(f^{-1}(\bar{\boldsymbol{X}}_\ell))$. Note that we can estimate $\bar{\boldsymbol{X}}_\ell$ using our data, so if we knew f^{U} and f, we could estimate $\bar{\boldsymbol{X}}_\ell^U$. The estimated $\bar{\boldsymbol{X}}_\ell^U$ would allow us to control for the effects of sorting in our regression. Altonji and Mansfield (2018) show that under a set of conditions that we will discuss in a moment, this mapping from observable averages to unobservable averages is in fact a *linear* function, i.e. $f^{\mathrm{U}}(f(X)) = XB$ for some invertible matrix B. Under thse assumptions, controlling for $\bar{\boldsymbol{X}}_\ell$ also controls for $\bar{\boldsymbol{X}}_\ell^{\mathrm{U}}$ which is the basis of our identificaiton strategy.

We have already listed three of the five assumptions required for this linear mapping result: (1) $\varepsilon_{i\ell}$ are assumed to be mean zero and drawn independently from observables, unobservables, and amenities of the locations; (2) preferences are given by equation 3; and (3) workers select their location taking prices and amenities as given.

The final two assumptions restrict the relationships between $\boldsymbol{X}_i$, $\boldsymbol{X}_i^{\mathrm{U}}$ and $\boldsymbol{A}_\ell$. First, we need the set of location-specific amenities that drive sorting on *observable* covariates contains *all* location-specific amenities that drive sorting based on *unobservable* covariates. As an example of where this might break down, say workers with high creative intelligence are sorting into cities based on the large amount of arts available. If no observable characteristics alters the worker's willingness to pay to live in an area with arts, then the 'inversion' of $\bar{\boldsymbol{X}}_i$ to $\boldsymbol{A}_\ell$ can not inform us about a location's arts amenity and therefore can not inform us about the average creative intelligence of it's workers.

Second, we require the mapping $f(\bar{\boldsymbol{X}}_\ell) = \boldsymbol{A}_\ell$ to be invertible. This requires (i) the vector of observable characteristics used must explain the willingness to pay for a wide range of amenities (f must be onto) and (ii) the number of observable averages we use must be larger than the number of amenities that workers chose from (f must be one-to-one). This is a daunting task since there might be a high-dimension of amenities that workers sort on.

However, in practice many amenities are 'bundled' together in the sense that a location high in one amenity is very likely to be high in all related ones.[14] Diamond (2016) provides evidence using a principal component analysis on a detailed set of amenities and documents that the effective rank of amenities is around five (with three components explaining most of the variation). In the context of this paper, the key identifying assumptions are that (i) our group-average variables have an effective (column) rank larger than five. Empirically, we run principal components analysis and find more than five components that explain more than 5 percent of the total variance across years. This provides support for using the approach in Altonji and Mansfield (2018) to address sorting.

Taking this approach, we rewrite equation (1) to include group-level averages of the individual characteristics in each location. In particular, we estimate:

$$\log \text{wage}_i = \beta_0 + \beta_1 \text{urban}_i + \boldsymbol{X}_i \delta + \bar{\boldsymbol{X}}_{c(i)} \gamma + \varepsilon_i, \tag{4}$$

where $\bar{\boldsymbol{X}}_{c(i)}$ are the location-specific average of the individual characteristics in $\boldsymbol{X}_i$ and additional location-invariant characteristics.[15] In this case, β_1 now captures the urban premium after controlling for sorting. Note that by controlling for the average characteristics of a worker's neighbors, we are, in effect, controlling for peer effects on a worker's wages. Our estimate may, therefore, underestimate the true impact of living in a city if it mainly operates through better peer-networks. As we will see below, given that we match other estimates found using panel data, this alleviates some concern.

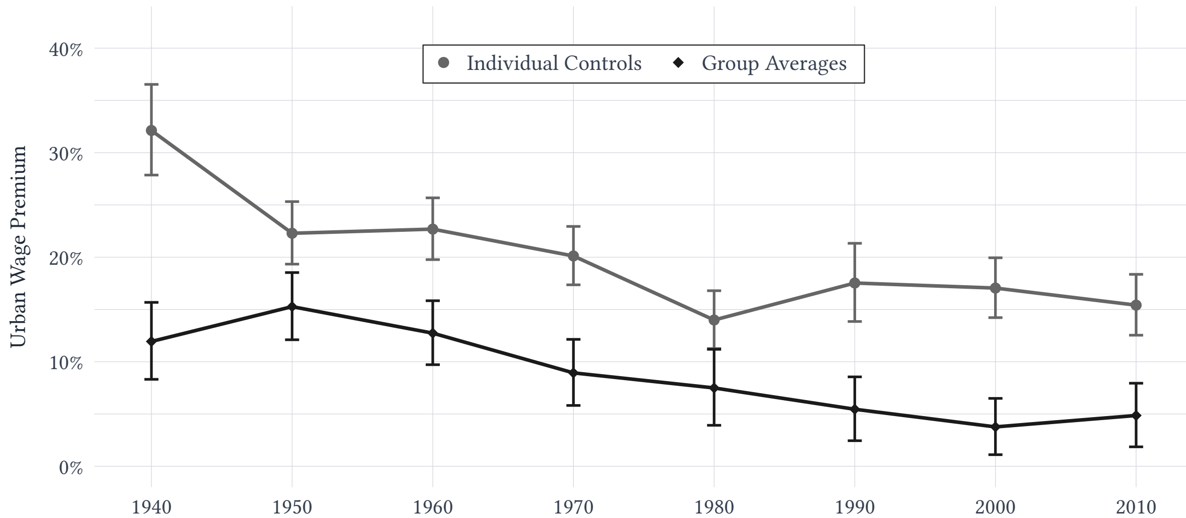
4.1 Estimated Causal Urban Wage Premium

We present our main estimates of equation (4) for each year are plotted in Figure 2. These estimates make clear that using the group-level averages of individual characteristics to control for sorting decreases the urban wage premium significantly. Relative to the raw urban

[14] For example, amenities such access to arts and access to dining are so highly correlated that workers are not able to 'shop for' one of them without receiving the other.

[15] We include as group-averages the share of white workers (the share of non-white workers drops out); share of residents in all but one of the age groups; share that are veterans; share that are married and together, married and apart, and single; and share with less than a high school degree, share with a high school degree, share with some college, share with BA degree, and share with advanced education. Additionally, we include a measure of market access calculated as in Jaworski and Kitchens (2019).

Figure 2: Urban Wage Premium Controlling for Individual Characteristics and Sorting Using Group-Level Averages, 1940-2010



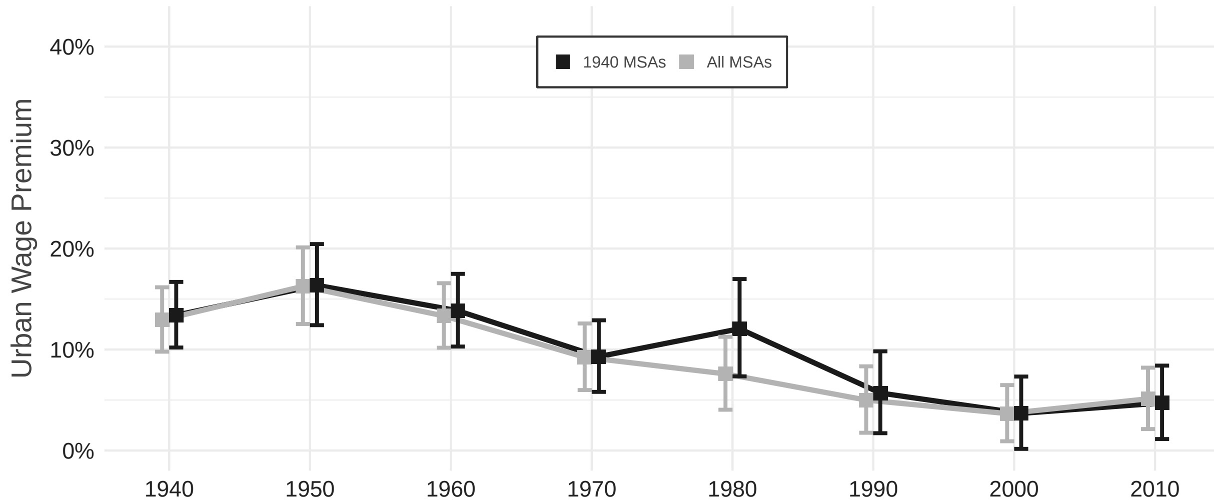
Notes: This figure plots the coefficient for survey-year level estimates of the urban wage premium following the methodology in Section 4. We estimate equation (4) with the individual-level and group-average controls. These point estimates serve as our estimate for the causal urban wage premium.

versus non-urban wage differential, our estimates suggest that individual characteristics and sorting explain approximately 75 percent of the wage differential between urban and non-urban areas after 1980 and approximately half in earlier decades.

Focusing on the period between 1980 and 2010, when our estimates overlap those reported in the existing literature, we find an urban wage premium of approximately 5 percent in each year. Our estimates closely match those reported in Glaeser and Maré (2001), who report a premium of 4.5 percent in the densest metropolitan areas. Our estimates also mirror those reported by Yankow (2006), whose fixed effects estimates range from 4.9 to 5.4 percent. Our estimates are also within the range of those reported in Baum-Snow and Pavan (2012). The overlap of our estimates for the decades since 1980 and previous estimates provides support for our approach using group-level averages to address the potential role of sorting.

Next, we turn to the period from 1940 to 1980, where the previous literature has yet to provide estimates of the urban wage premium. In 1940, we find that the urban wage premium was approximately 13 percent. Given a raw differential of 31 percent, sorting and

Figure 3: Urban Wage Premium Using the 1940 Definition of MSAs, 1940-2010



Notes: This figure plots the coefficient for survey-year level estimates of the urban wage premium following the methodology in Section 4. We estimate equation (4) with the individual-level and group-average controls using only the subsample of 1940 MSAs. Individuals living in MSAs that were added after 1940 are dropped from the sample. The estimates from Figure 2 are reproduced for comparison.

individual characteristics explain about three-fifths of the urban versus non-urban wage gap. Between 1940 and 1970, we estimate that the urban wage premium fell substantially. The premium stayed high at 16 percent in 1950 and 13 percent in 1960 and then began to decline to about 5 percent starting in the 1990s.

The estimates for the period from 1940 to 1970 are striking. Our estimates for the earlier period are more than twice the size of those reported in the literature for the period between 1980 and 2020. The recent literature has sought to explain the widening urban versus non-urban wage gap since 1970, largely focusing on the differential return to skill in urban areas (Ganong and Shoag, 2017; Baum-Snow, Freedman and Pavan, 2018; Dauth, Findeisen, Moretti and Suedekum, 2018; Autor, 2019). For example, Baum-Snow et al. (2018) and Autor (2019) highlight that the returns to education in urban areas increased starting in 2000 due to agglomeration forces that are skill-biased.

Our estimates suggest that earlier periods of rapid change affected the wage differential between urban and non-urban areas. We show that sorting became increasingly important in postwar period, which is earlier than previously documented. Broadly, these results reinforce the focus of Autor (2019) and Papageorgiou (2022) on the changing number and composition of occupations available for low-skill workers in this period. For example, our estimates are consistent with the wave of factory automation in the 1950s and early 1960s.

Over time, the number of MSAs grows as new cities increase in size. This raises a concern that our estimates may reflect a change in the composition of MSAs. To rule this out, we re-estimate equation (4) with a fixed sample of MSAs based on their 1940 designations and drop observations from MSAs designated after 1940 from the sample. In Figure 3 we report the estimates holding the set of MSAs constant. The results are similar to our baseline estimates, suggesting that the changing composition of cities does not bias our estimates.

4.2 Causal Urban Wage Premium Adjusted for Housing Costs

So far we have ignored the role that rents have in determining an individual's 'real wage'. Canonical urban models suggest that urban rents will adjust in response to higher wages and partially absorb the wage premium, such that utility is equalized across space.

To ensure that the urban wage premium we estimate is not driven by rent differentials, we re-estimate equation (4) using a measure of income that accounts for location-specific differences in housing costs.[16] Following Ganong and Shoag (2017) and Hoxie et al. (2020), we use both rent and home value in calculating the cost of housing. For rent, we use the monthly rent and divide it by 4 to get weekly rent.[17] For home values, we set the annual housing cost at 5% of the home value and divide by 48 for weekly costs. Then our housing cost is subtracted from the reported weekly wage to get that individual's net-wage.

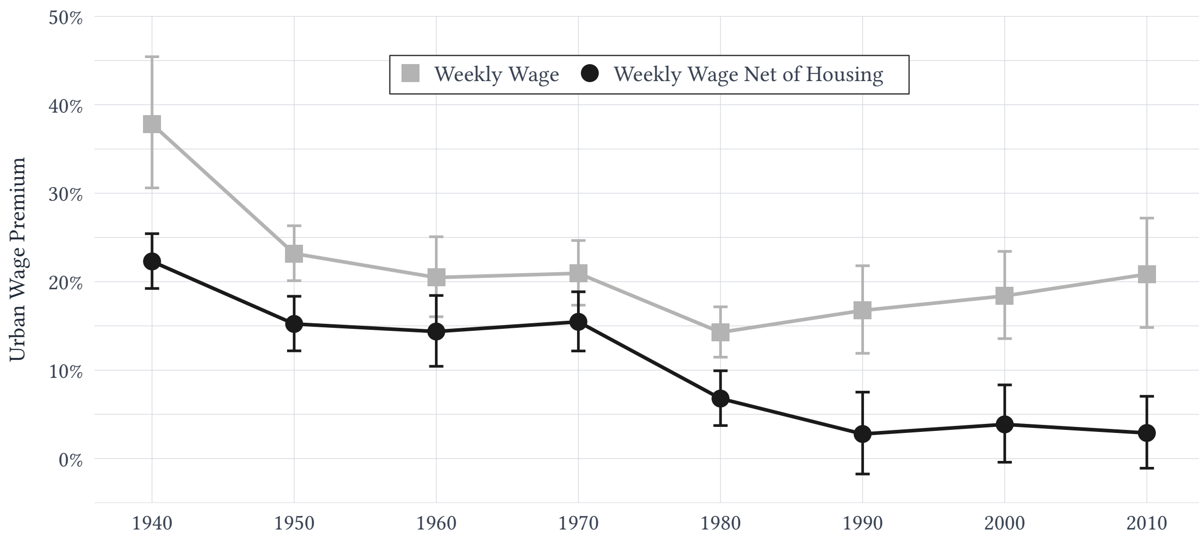
Figure 4A plots the raw wage premium after subtracting the cost of housing. This figure shows that in the raw data, a portion of the wage premium is extracted by landlords via a higher rent premium. Turning to our causal estimates in figure 4B, the results after 1970 do not seem to significantly change. However, the estimated wage premium from 1940-1960 shrinks slightly to between 5 and 10%. These results suggest that some of the causal benefits

[16] For example, between 1960 and 2010, the average rent in an MSA was between 50 and 70 percent higher than non-MSA regions.

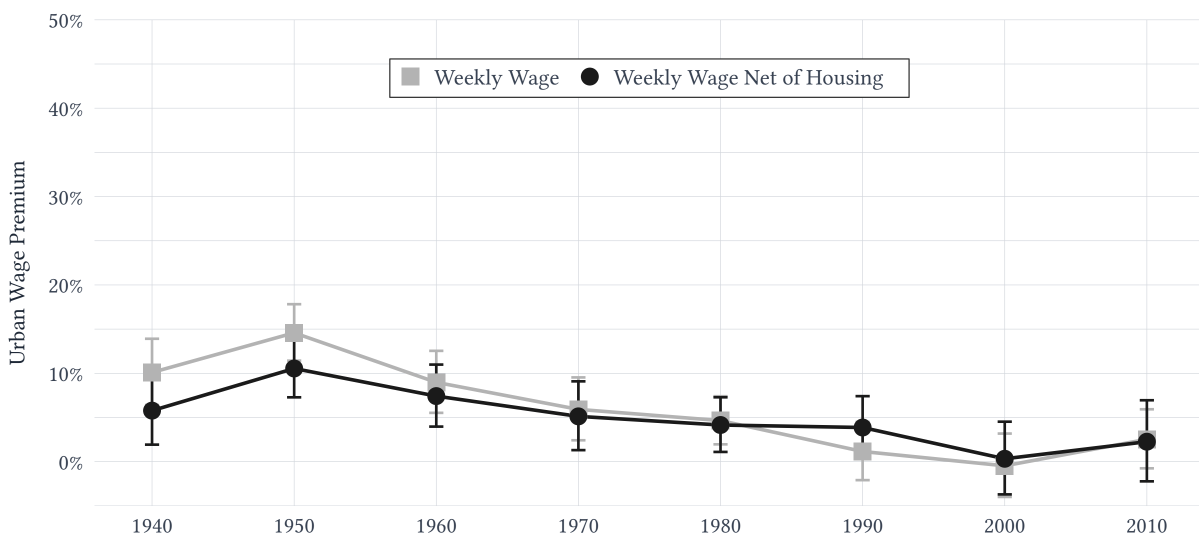
[17] For 1950, microdata on rent is not available, so instead we use the average rent as reported in census summary tables.

Figure 4: Urban Wage Premium Net of Housing Costs, 1940-2010

(A) Raw Differences



(B) Controlling for Individual Characteristics and Sorting Using Group-Level Averages



Notes: This figure recreates figures 1 and 2 using a modified wage measure that subtracts off the cost of housing.

of cities that accrued to worker was lost to higher rents between 1940 and 1960.

5 Heterogeneity by City Size, Region, and Indvidiaul Characteristics

Up to this point, we have documented the increasing role of sorting in explaining the difference between urban and non-urban wages over time. Importantly, this pattern may differ substantially throughout the city-size distribution, across regions, or for workers of different characteristics. In this section, we examine the role of these factors.

Recent literature highlights that the urban wage premium is increasing in city size (e.g. Baum-Snow and Pavan, 2012). To test for heterogeneity in the urban wage premium throughout the city-size distribution, we re-estimate equation (4) but allow cities in the 20 most populated MSAs as of 1940 to have a different urban wage premium. Panel A of Figure 5 plots the estimated urban wage premium for the twenty largest MSAs (square markers) versus the remaining MSAs (round markers). First, we document that the urban wage premium is higher in the largest 20 MSAs relative to smaller MSAs in all years, consistent with the existing literature. While the largest MSAs generate larger wage premia, the returns to the largest MSAs have declined relative to smaller MSAs in the last 30 years.

Previously, we highlighted how the geography of the highest-wage cities has varied over time. In the 1940s, the largest premiums were in Rust Belt cities such as Flint, MI and Rochester, NY. Today, the pattern has shifted towards Sunbelt cities and technology hubs. Given the shift in the location of work, the relative importance of sorting and agglomeration may also differ across space over the last 80 years. To explore how the relative importance of sorting varies across geography, we estimate specifications to allow for differences in each census region (i.e., Northeast, South, Midwest, and West). Panel B of Figure 5 plots the estimates. The Midwest has experienced a large decline in the share of the premium attributable to agglomeration, while agglomeration forces have become relatively more important in the West since 1970.

As automation and technology have increasingly entered the workplace, the returns to cities may also vary over time for workers with different levels of education. For example, Gould (2007) shows that the urban wage premium is significantly larger for white-collar workers. To explore how agglomeration forces have varied by education group, we estimate separate specifications for college-educated workers and those with less than a college degree.[18] Panel C of Figure 5 plots our estimates by college degree status. Throughout the period, we generally estimate larger wage premia for college-educated workers, on the order of magnitude of 3-5 percentage points. Comparing our estimates to Gould (2007), we similarly find a premium of around 10 percent for college-educated workers in the 1980s.

We run 1 while interacting the urban dummy with a dummy for being a White worker and a dummy for being a Black worker. This estimates a race-specific urban wage premium. Panel C of Figure 5 shows the results. Throughout the panel, the Black urban wage premium is estimated to be no smaller than the White premium and in most of the years a larger. The Black urban wage premium is estimated to be ten percentage points larger in the middle of the 20th century which is in line with the timing of the Great Migration (Derenoncourt, 2022).

Last, Baum-Snow and Pavan (2012) document the urban *wage growth* premium which shows that cities benefit workers by faster wage growth. Since we do not have panel data, we can not observe wage growth. Instead, we run 1 separately for indvidiauls in above-median and below-median aged workers[19] The idea being that older workers have more *potential* time in city and therefore older workers that live in urban areas likely have more cummulative urban experience than younger workers and hence should have a larger premium. This hypothesis is borne out by Panel E of Figure 5. An alternative explanation is that the benefits of urban living change over the workers' earning profile.

5.1 Distributional Estimates

In the period after World War II, there were substantial changes in wage inequality. Goldin and Margo (1992) discuss the narrowing of the wage distribution in what they label the "Great Compression" between 1940 and 1960. They highlight the rapid increase in the supply increase of educated workers during a period of relatively stagnant demand growth. Recently, Vickers and Ziebarth (2021) discuss the role of the National War Labor Board on

[18] For panels C, D, and E of Figure 5, we calculate the group averages separately by group since we may think that sorting patterns can differ systematically.

[19] The median age is around 40 across years.

wages in various regions throughout the country and show that wage-setting policies had a meaningful impact on local wage distributions.

To explore whether we observe a compression and then widening of the wage distribution in our setting, we take the residuals from (4) which reflect individual earnings after controlling for sorting. The ninetieth-tenth percentiles summarize the distribution of the wage residuals for each decade in Panel A of Figure 6. Similar to Goldin and Margo (1992), we find evidence consistent with a narrowing of the wage distribution through 1960. Afterward, the wage distribution widens considerably through 2010.

In Panel B of Figure 6, we plot the distribution of residuals for urban and non-urban workers. In 1940, urban and non-urban workers had a large left tail in the wage distribution, reflecting many workers who earned very low (residualized) weekly earnings. Between 1940 and 1960 this left-tail disappeared, which provides some evidence that a rising floor potentially drove the Great Compression on non-urban wages.[20] Goldin and Margo (1992) find similar evidence that increases in the distribution for low-wage workers drove the Great Compression. Our estimates suggest that the narrowing of the wage distribution occurred in urban and non-urban areas.

6 Conclusion

Wages in cities are much higher than in non-urban areas. Understanding why this gap exists and, in particular, the relative role of sorting versus agglomeration over time is an important topic in urban economics. In this paper, we quantify the urban wage premium in each decade between 1940 to 2010. Importantly, we show how to apply the methodology of Altonji and Mansfield (2018), which enables us to use data from annual cross-sections to account for sorting between urban and non-urban locations. Thus, we provide new estimates of the urban wage premium for the period from 1940 to 1980, while generating estimates for the period from 1980 to 2010 that are similar to those in the existing literature.

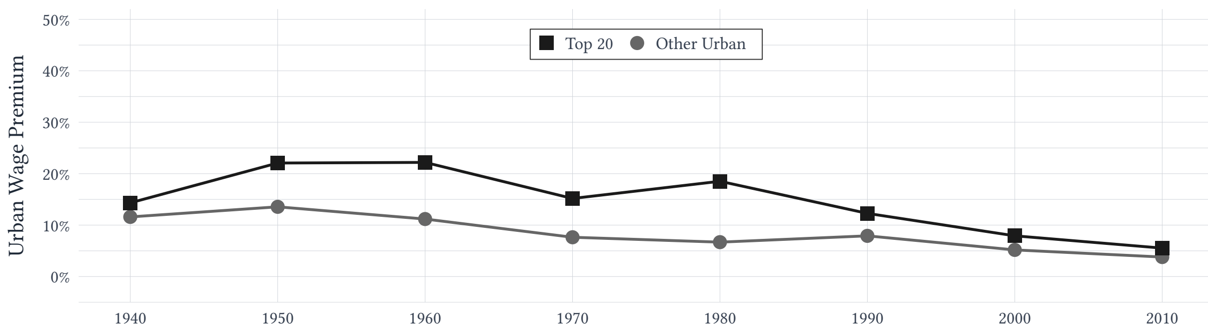
Our results highlight the increasing relative importance of sorting over the second half of the twentieth century and into the first part of the twenty-first century. By 2010, the relative

[20] The overall compression and the narrowing of the left-tail are also apparent when using 95^{th} and 5^{th} percentiles.

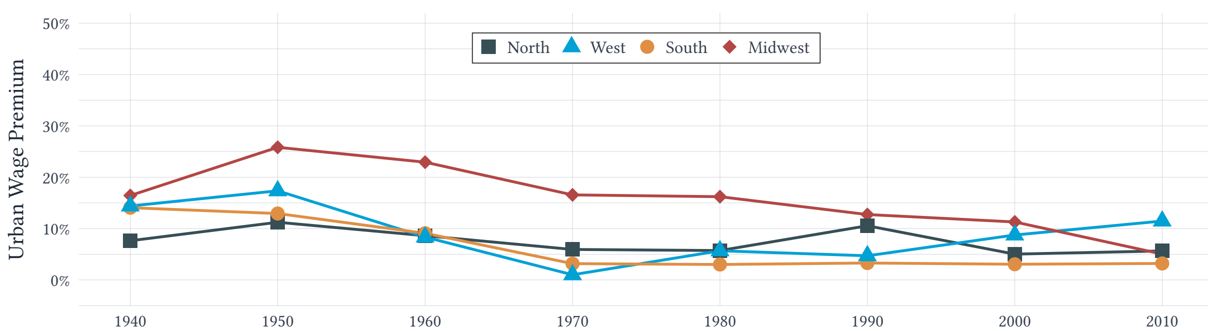
strength of agglomeration forces was less than half of the agglomeration forces we estimate in 1940. In addition, we find that this pattern is not driven by changes in the composition of cities in our sample over time or differences in housing costs across cities. We also show that changes in the urban wage premium and the role of sorting are not due to changes in the returns to city size, regions, level of education, or the structure of the wage distribution across urban and non-urban areas. Overall, our findings are consistent with a greater role for consumer amenities in explaining the demand for urban living as of 2010.

Figure 5: Heterogeneity of the Urban Wage Premium, 1940-2010

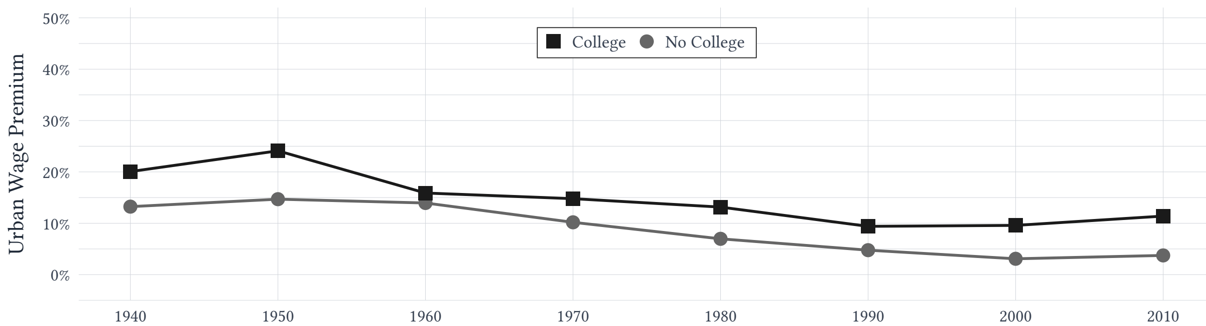
(A) By Twenty Largest versus Other Cities



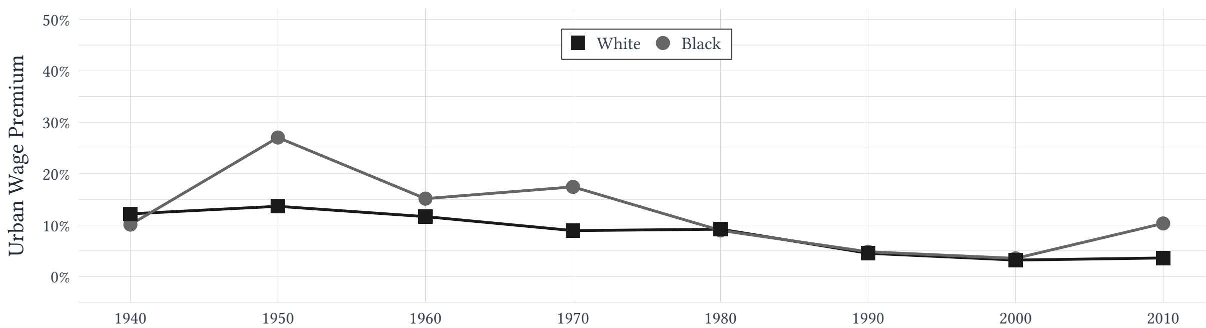
(B) By Census Regions



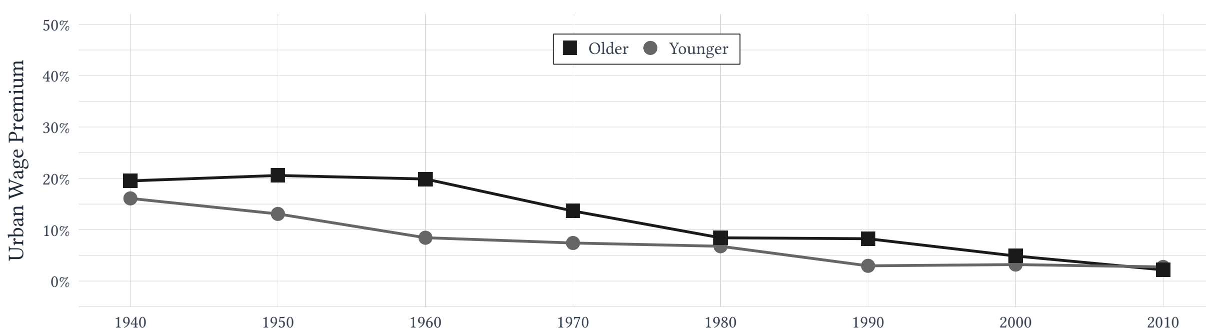
(C) College and Non-college Educated Workers



(D) White and Black Workers



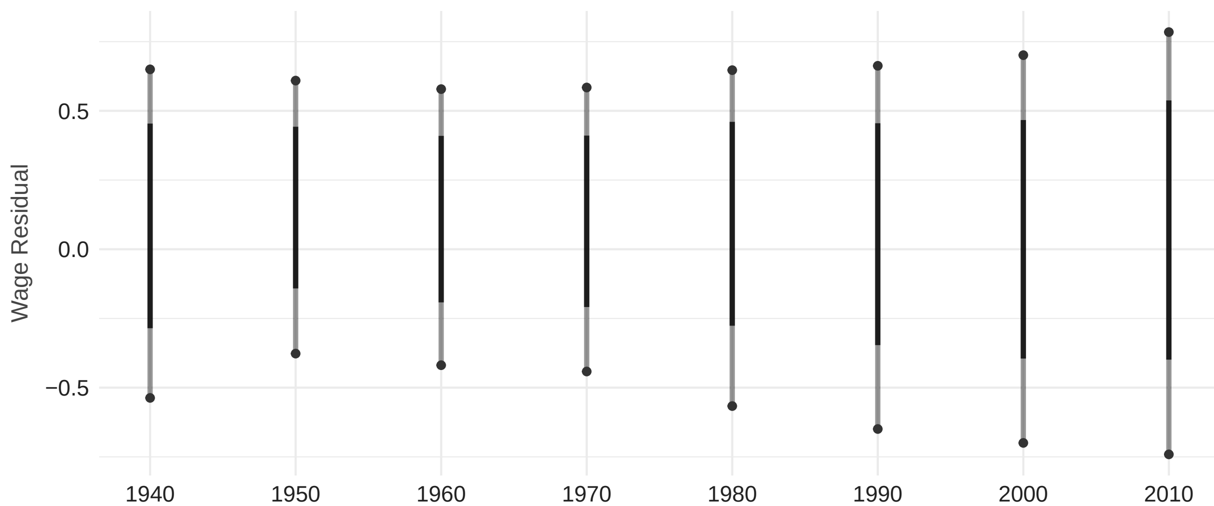
(E) Older and Younger Workers



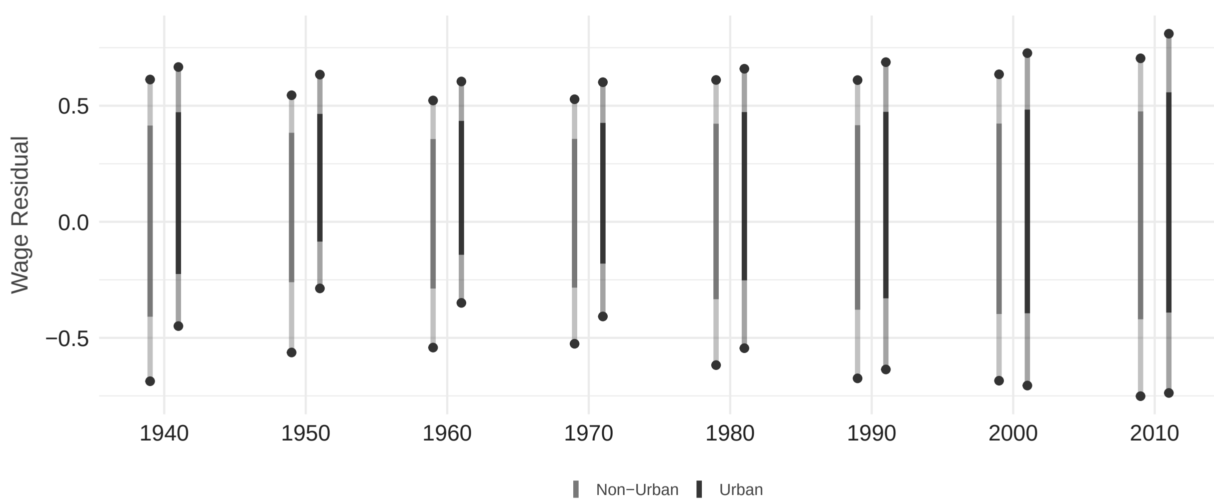
Notes: This figure plots the coefficient for survey-year level estimates of the urban wage premium from equation (1) including the individual-level and group-average controls described in footnote 15 while interacting the urban indicator with three different variables. Panel A interacts the urban indicator with an indicator for the top 20 most populous MSA in 1940. Panel B interacts the urban indicator with indicators for each census region. Panel C reports estimates of equation (1) run on separate samples of college-educated and non-college-educated workers with the group-averages taken separately by sample.

Figure 6: Distribution of Wage Residuals, 1940-2020

(A) Aggregate Distribution



(B) Urban versus Non-urban Distribution



Notes: This figure takes the residual of wages after regressing on the full-specification and adds back in the causal urban wage premium. Each panel plots points for the 90th-10th percentiles of the wage residuals and shades darker the 80th-20th percentiles. Panel A does this for the entire distribution of wages following Goldin and Margo (1992). Panel B does this separately for urban and non-urban workers.

References

Ahlfeldt, Gabriel M, Thilo Albers, and Kristian Behrens, "A granular spatial model," Technical Report 2022.

Akcigit, Ufuk, Salomé Baslandze, and Stefanie Stantcheva, "Taxation and the International Mobility of Inventors," *American Economic Review*, 2016, *106* (10), 2930–81.

Altonji, Joseph G. and Richard K. Mansfield, "Estimating Group Effects Using Averages of Observables to Control for Sorting on Unobservables: School and Neighborhood Effects," *American Economic Review*, 2018, *108* (10), 2902–46.

Autor, David H., "Work of the Past, Work of the Future," *AEA Papers and Proceedings*, May 2019, *109*, 1–32.

Baum-Snow, Nathaniel and Daniel Hartley, "Accounting for Central Neighborhood Change, 1980–2010," *Journal of Urban Economics*, 2020, *117*, 103228.

_ and Ronni Pavan, "Understanding the City Size Wage Gap," *Review of Economic Studies*, 2012, *79* (1), 88–127.

_ , Matthew Freedman, and Ronni Pavan, "Why Has Urban Inequality Increased?," *American Economic Journal: Applied Economics*, 2018, *10* (4), 1–42.

Behrens, Kristian, Gilles Duranton, and Frédéric Robert-Nicoud, "Productive Cities: Sorting, Selection, and Agglomeration," *Journal of Political Economy*, 2014, *122* (3), 507–553.

Boustan, Leah Platt, Devin Michelle Bunten, and Owen Hearey, "Urbanization in the United States, 1800-2000," Working Paper 19041, National Bureau of Economic Research May 2013.

Combes, Pierre-Philippe, Gilles Duranton, and Laurent Gobillon, "Spatial Wage Disparities: Sorting Matters!," *Journal of Urban Economics*, 2008, *63* (2), 723–742.

Dauth, Wolfgang, Sebastian Findeisen, Enrico Moretti, and Jens Suedekum, "Matching in Cities," *Journal of the European Economic Association*, 2018.

_ , _ , _ , and _ , "Matching in cities," *Journal of the European Economic Association*, 2022, *20* (4), 1478–1521.

Derenoncourt, Ellora, "Can you move to opportunity? Evidence from the Great Migration," *American Economic Review*, 2022, *112* (2), 369–408.

Diamond, Rebecca, "The Determinants and Welfare Implications of US Workers' Diverging Location Choices by Skill: 1980–2000," *American Economic Review*, Mar 2016, *106* (3), 479–524.

_ and Cecile Gaubert, "Spatial Sorting and Inequality," *Working Paper*, 2021.

Dingel, Jonathan I and Felix Tintelnot, "Spatial economics for granular settings," Technical

Report, National Bureau of Economic Research 2020.

Dumais, Guy, Glenn Ellison, and Edward L. Glaeser, "Geographic Concentration as a Dynamic Process," *Review of Economics and Statistics*, 2002, *84* (2), 193–204.

Eckert, Fabian, Mads Hejlesen, and Conor Walsh, "The return to big-city experience: Evidence from refugees in Denmark," *Journal of Urban Economics*, 2022, *130*, 103454.

Frings, Hanna and Rebecca Kamb, "The Relative Importance of Portable and Non-Portable Agglomeration Effects for the Urban Wage Premium," *Regional Science and Urban Economics*, 2022, *95*, 103786.

Ganong, Peter and Daniel Shoag, "Why Has Regional Income Convergence in the US Eeclined?," *Journal of Urban Economics*, 2017, *102*, 76–90.

Glaeser, Edward L. and David C. Maré, "Cities and Skills," *Journal of Labor Economics*, Apr 2001, *19* (2), 316–342.

_ and Giacomo A.M. Ponzetto, "Did the Death of Distance Hurt Detroit and Help New York?," *Agglomeration Economics*, 2010, *303*, 303–05.

_ and Jesse M. Shapiro, "Urban Growth in the 1990s: Is City Living Back?," *Journal of Regional Science*, 2003, *43* (1), 139–165.

_ and Joshua D. Gottlieb, "The Wealth of Cities: Agglomeration Economies and Spatial Equilibrium in the United States," *Journal of Economic Literature*, 2009, *47* (4), 983–1028.

_ , Jed Kolko, and Albert Saiz, "Consumer City," *Journal of Economic Geography*, 2001, *1* (1), 27–50.

Goldin, Claudia and Robert A Margo, "The great compression: The wage structure in the United States at mid-century," *The Quarterly Journal of Economics*, 1992, *107* (1), 1–34.

Gould, Eric D., "Cities, Workers, and Wages: A Structural Analysis of the Urban Wage Premium," *Review of Economic Studies*, 2007, *74* (2), 477–506.

Hoxie, Philip, Daniel Shoag, and Stan Veuger, "Moving to Density: Half a Century of Housing Costs and Wage Premia from Queens to King Salmon," Technical Report, AEI Economics Working Paper 2020.

Jaworski, Taylor and Carl T. Kitchens, "National Policy for Regional Development: Historical Evidence from Appalachian Highways," *Review of Economics and Statistics*, 2019, *101* (5), 777–790.

Kerr, William R and Frédéric Robert-Nicoud, "Tech clusters," *Journal of Economic Perspectives*, 2020, *34* (3), 50–76.

Kline, Patrick and Enrico Moretti, "People, Places, and Public Policy: Some Simple Welfare Economics of Local Economic Development Programs," *Annual Review of Economics*, Aug 2014, *6* (1), 629–662.

Martin, Philippe and Gianmarco IP Ottaviano, “Growth and agglomeration,” *International economic review*, 2001, *42* (4), 947–968.

Moretti, Enrico, “Local Labor Markets,” in “Handbook of Labor Economics,” Vol. 4 2011, pp. 1237–1313.

_ , “Real Wage Inequality,” *American Economic Journal: Applied Economics*, 2013, *5* (1), 65–103.

_ and Daniel J. Wilson, “The Effect of State Taxes on the Geographical Location of Top Earners: Evidence from Star Scientists,” *American Economic Review*, 2017, *107* (7), 1858–1903.

Papageorgiou, Theodore, “Occupational matching and cities,” *American Economic Journal: Macroeconomics*, 2022, *14* (3), 82–132.

Puga, Diego, “The Magnitude and Causes of Agglomeration Economies,” *Journal of Regional Science*, 2010, *50* (1), 203–219.

Rosenthal, Stuart S. and William C. Strange, “Evidence on the Nature and Sources of Agglomeration Economies,” in “Handbook of Regional and Urban Economics,” Vol. 4, Elsevier, 2004, pp. 2119–2171.

Ruggles, Steven, Catherine A. Fitch, Ronald Goeken, J. David Hacker, Matt A. Nelson, Evan Roberts, Megan Schouweiler, and Matthew Sobek, “IPUMS Ancestry Full Count Data: Version 3.0 [dataset],” in “,” IPUMS, 2022.

_ , Sarah Flood, Ronald Goeken, Megan Schouweiler, and Matthew Sobek, “IPUMS USA: Version 12.0 [dataset],” in “,” IPUMS, 2022.

Schoefer, Benjamin and Oren Ziv, “Productivity, Place, and Plants,” *Review of Economics and Statistics*, 2022, pp. 1–46.

Shapiro, Jesse M., “Smart Cities: Quality of Life, Productivity, and the Growth Effects of Human Capital,” *Review of Economics and Statistics*, 2006, *88* (2), 324–335.

Su, Yichen, “The Rising Value of Time and the Origin of Urban Gentrification,” *American Economic Journal: Economic Policy*, 2022, *14* (1), 402–39.

Verginer, Luca and Massimo Riccaboni, “Talent Goes to Global Cities: The world Network of Scientists’ Mobility,” *Research policy*, 2021, *50* (1), 104127.

Vickers, Chris and Nicolas L Ziebarth, “The Effects of the National War Labor Board on Labor Income Inequality,” *Working Paper*, 2021.

Yankow, Jeffrey, “Why Do Cities Pay More? An Empirical Examination of Some Competing Theories of the Urban Wage Premium,” *Journal of Urban Economics*, 2006, *60* (2), 139–161.

Table 3: Urban Wage Premium, 1940-2010

	log(Weekly Wage)							
	1940	1950	1960	1970	1980	1990	2000	2010
Panel A: Raw Wage Premium								
Urban = 1	0.3162***	0.2208***	0.2288***	0.2109***	0.1692***	0.1939***	0.2014***	0.1898***
	(0.0192)	(0.0131)	(0.0134)	(0.0119)	(0.0110)	(0.0158)	(0.0144)	(0.0155)
Panel B: Individual Controls								
Urban = 1	0.2822***	0.2012***	0.2038***	0.1833***	0.1334***	0.1609***	0.1566***	0.1414***
	(0.0186)	(0.0125)	(0.0123)	(0.0119)	(0.0116)	(0.0163)	(0.0125)	(0.0127)
Panel C: Individual Controls & Group Averages								
Urban = 1	0.1216***	0.1506***	0.1251***	0.0883***	0.0731***	0.0487***	0.0359***	0.0499***
	(0.0144)	(0.0167)	(0.0144)	(0.0154)	(0.0171)	(0.0160)	(0.0137)	(0.0147)
Observations	2,504,957	68,457	1,364,524	280,338	1,907,943	2,260,990	2,582,144	2,718,970

Notes. This table reports coefficient for survey-year level estimates of the urban wage premium from equation (1) with three different level of controls. The first section includes no controls and represents the urban wage premium observed in the data. The second section includes controls for age in 5-year bins, education levels, and an indicator for being White. The final section includes group-average variables following the methodology in Section 4. These point estimates serve as our estimate for the causal urban wage premium.
Signif. Codes: ***: 0.01, **: 0.05, *: 0.1. Standard errors are clustered by MSA/CBSA.