The Evolution of the Spatial Wage Inequality Across the Work-Cycle in the U.S.

Pablo Andrés Valenzuela Casasempere *

University of British Columbia

October 1, 2021

Abstract

This paper documents the heterogeneous rise in the urban gradient of the college wage gap across workers of different ages between 1980 and 2019. Using immigrants' enclaves from 1970 as source of identification, I find that the young workers have traditionally had a steeper relationship between college wage gap and city population than old workers. Also, I find that the evolution of this urban gradient of the college wage gap has been larger for younger workers. These findings are not caused by sorting in unobserved characteristics, by outliers in the wage distribution, or by compositional changes. I show that the source of the increase in the urban component of the college wage gap is a shift in the occupational structure across the work-cycle and cities. While old and young college workers have shifted away from highly-routinary-low-paying jobs, specially in more populated cities, young high school graduates' occupational structure has remained unaltered since 1980.

*pablo.valenzuela@ubc.ca

1 Introduction

In the last forty years the U.S. labour market has been characterized by a significant increase in the average wage gap between high school and college graduates. This gap is not population neutral. College and high school workers face larger differences between their hourly wages in more populous cities. This relationship is referred to as the college bias of agglomeration externalities. The college bias of agglomeration externalities has only increased over time. For example, in 1980 the (log) college wage gap was 0.05 larger in the most versus the least populated city. In 2019 this difference was over 0.20. Previous work studying the college bias of agglomeration externalities, however, falls short of providing a satisfactory analysis of the effect agglomeration externalities have on the work-cycle. The literature has assumed that workers of different ages are perfect substitutes and, consequently, are impacted homogeneously by agglomeration externalities. Nonetheless, this assumption doesn't hold in the data.

In this paper I relax the assumption of perfect substitution across workers of different ages. I show that the shifting structure of the returns to college education in larger U.S. cities is a reflection of an increase in the age-group specific college bias of agglomeration externalities, which attracts young college educated workers to larger cities. As a result, the relative supply of college educated workers in more populous cities has increased with each passing decade. I find that a larger age-specific relative supply of college workers is associated with a decrease in the college wage gap. However, this decrease is smaller in magnitude than the increase caused by the rise in the college bias of agglomeration externalities. Thus, explaining why the college wage gap has increased in more populous cities.

This paper provides estimates of the college bias of agglomeration externalities by different age groups. This exercise is of interest for two reasons. First, it furthers our understanding on how agglomeration externalities affect workers of different educational attainment and how it varies across different stages of workers' careers. In the last decades, bigger cities have become younger and more college intensive as the increase in the college wage gap for young workers was unmatched by any other age group (Lee et al., 2019). Therefore, this project extends and complements the economic literature studying the spatial component of the college wage gap.

Second, it deepens our knowledge on the geography of labour inequality. To interpret the pronounced increase of the college wage premia in the United States it is necessary to understand the rise in the agglomeration externalities across cities and workers' work-cycle. Baum-Snow and Pavan (2013) show that at least one-quarter of the increase in nationwide wage inequality since 1980 can be attributed to faster increases in the college wage premia in larger cities. This project complements this literature by analyzing how wage inequality across cities is distributed through the work-cycle and its interaction with agglomeration externalities.

I propose a theoretical framework to quantify the factors that impact the college wage across cities of different populations. In particular, I employ a model of city aggregate production with age-group specific labour supplies that incorporates imperfect substitution between workers of different age and educational attainment, building upon Card and Lemieux (2001) and Baum-Snow et al. (2018) models of nested constant elasticity of substitution. The model allows for heterogeneous effect of agglomeration externalities across the work-cycle and between different educational levels. The model predicts that the college wage gap in a given city will increase: (i) if the age-specific relative supply of college workers decrease, (ii) if the overall relative supply decrease, (iii) if there is a technological shock that benefits college over high school workers, or (iv) if agglomeration externalities benefit more college educated workers than high school workers.

Next, I estimate the predictions done by the theoretical model using five decades of U.S. census data. The results indicate that changes in the college bias of agglomeration externalities have been central to the rise of wage inequality across local labour markets. I find that for all workers the college bias of agglomeration externalities has steepened over time. However, this effect has been particularly large for workers under forty years old. The decadal increase in the college wage gap for young workers located in the city with the largest population is 3.7% and 7.1% larger than the increase in the median city in 1990 and 2019. For old workers the same effect is 2.4% and 3.4%, respectively. The heterogeneous evolution in the college bias of agglomeration externalities is not caused by sorting on unobserved characteristics, by outliers in the wage distribution, or by compositional effects.

The empirical strategy used allows me to recover the elasticity of substitution between workers of different ages and between workers with different educational attainment. I find an elasticity of substitution between age groups of 4.36, which supports the assumption of imperfect substitution between workers of different ages. Although my goal is to measure the college bias of agglomeration externalities, my estimates of the elasticity of substitution between age groups is comparable in magnitude to those in Card and Lemieux (2001). I also estimate an elasticity of substitution between workers with different educational attainment of 3.6. This result is larger to the results in Katz and Murphy (1992), Bound and Johnson (1992), and Autor et al. (2008), who find an elasticity around 1.4. The main reason of this difference is the decades used in the sample. Since the college premia has increased over time, studies which include larger periods tend to find a larger elasticity of substitution between college and high school workers. My estimates, for example, are close in magnitude to new studies using larger time spans such as Baum-Snow et al. (2018). The fact that the estimates for both elasticities are in line with previous estimates from the literature can be taken as evidence in support of my model.

I find that the shift in the occupational structure across the work-cycle and cities explains the increase in the college bias of agglomeration externalities. While old workers have shifted away from highly routinary jobs regardless of their educational attainment or the city they live in, the same is not true for young workers. The share of young workers with a high school diploma performing highly routinary jobs in 2019 has the same level and slope with respect to city size that it had in 1980. Conversely, with each passing decade young workers with a college degree have reduced their share and slope with respect to city population in highly routinary jobs while simultaneously increasing their share in high-paying low routinary jobs. Thus, the heterogeneous change in the occupational structure explains the increase in the college bias of agglomeration externalities.

The contributions of this project spans into three literatures. First, the paper relates to the literature that explores local determinants of the U.S. labour market. Moretti (2013), Lindley and Machin (2014), and Baum-Snow and Pavan (2013) argue about the importance of geography in explaining recent trends in the American labour market inequality. Complementing those findings, Baum-Snow et al. (2018) and Giannone (2017) find that skills, proxied by the workers' educational attainment, and cities have become more complementary over time –agglomeration externalities for college workers have risen over time— explaining the increase in college wage gap in larger cities. Autor (2019) and Eckert et al. (2019) give two different explanations for this. Autor (2019) argues that the occupations performed by non-college workers in urban labour markets have changed in the last decades. In other words, non-college workers now perform essentially the same jobs in urban and non-urban labour markets. Eckert et al. (2019) argue that the rise in wages in industries that heavily rely on ICT adoption and college workers is the leading cause of this increase. In addition to corroborating the increasing complementary between skills and cities, the main contribution of the paper is to document and analyze the heterogeneous trajectory this complementarity has across the work-cycle. To the best of my knowledge, this is the first paper that documents these heterogenous impacts.

Second, this paper also relates to the literature looking at the effect agglomeration externalities have on the labour market (Glaeser and Mare, 2001; Wheeler, 2006; Combes and Gobillon, 2015; Combes et al., 2008). Higher wages in more populated places can be caused by (i) the sorting of workers with higher unobserved ability into larger cities, (ii) a static wage premium for working in larger places, and (iii) a higher premium for work experience acquired in larger places. This paper complements this body of literature by analyzing the dynamics of the city size wage premium for college and high school graduates over time and across the work-cycle. De La Roca and Puga (2017) find that places with a higher population have higher wage premiums and that experience gained in larger cities is more valuable. Similarly, I find that the city size premium is positive for college workers. For non-college workers, however, I find that the city size premium decreased over each passing decade, in line with the results of Gould (2007). This decrease is specially prevalent in workers at early stages of their careers.

Finally, this paper relates to the literature exploring the change in the occupational structure of the U.S. labour market and its impact on wage inequality (Acemoglu and Restrepo, 2021; Autor, 2019; Duranton and Puga, 2019; Davis et al., 2020; Lindley and Machin, 2014). Autor et al. (2003) note that rapid computerization alters the demand for jobs and the occupational structure of the labour force. Complementing these results, Autor and Dorn (2013) estimate that the wage and employment share growth of highly routinary jobs – those easily replaceable by computers – was lower than traditionally less skilled jobs. My contribution to this literature is to document the geographical variation in the occupations performed by workers at different ages, and its relationship with the increase in the college bias of agglomeration externalities.

In the next section I describe the data sources. Section 3 provides a descriptive examination of the geographical changes in wage inequality since 1980. In Section 4 I outline the model of imperfect substitution between workers of different educational attainment and age, match the theoretical model to the empirical equation I take to the data, and discuss the identification strategy. Section 5 presents the main results of the paper and in Section 6 I show a battery of robustness checks to these results. Section 7 studies the underlying mechanisms behind the results and finally Section 8 concludes.

2 Data

Large samples are essential for the analysis of changes in local labour market outcomes at the city level. To this end, I draw on the Census Integrated Public Use Micro Samples for the years 1970, 1980, 1990, and 2000, and the American Community Survey (ACS) 5-year sample for the years 2010 and 2019 (Ruggles et al., 2020).[1] The 1970 census sample corresponds to 1 percent of the population. The other five samples correspond to 5 percent of the US population.

The workers sample consists in male individuals between the ages of 26 and 60 who worked at least 40 hours a week for a minimum of 50 weeks a year. I excluded workers in the military force because they are not ruled by local labour forces. Labour supply is measured by the product of the usual hours worked times the total number of weeks in the year.[2] College workers are those with a college degree or more whereas high school workers correspond to the complementary group. Hourly wages are constructed by dividing annual wage and salary earnings by the the individual labour supply. I drop individuals with an hourly wage below 75% of the federal minimum wage in the sample

[1] The two samples drawn from the ACS correspond to 1% ACS samples for the five years before. For example, the 2019 ACS 5 year sample contains information from the 2015, 2016, 2017, 2018, and 2019 1% ACS samples.

[2] The main reason I use full time full year workers is because 2010 and 2019 ACS do not record the exact number of weeks worked. Those datasets report binned values of the variable. As it can be seen in Figure B.1, the majority of workers in the 50+ category work 52 weeks in a year. Thus, I impute 52 weeks as the weeks worked for all the years in my sample to reduce the imputation error.

year and those whose income, usual hours worked, or weeks worked are top coded.[3] All calculations are weighted by the census sampling weight multiplied by the labour supply weight and a weight derived from the geographic matching process that is described below.

An important feature of the data used is that its based on differences between individuals *of the same age* with and without a college degree. For example, the wage gap is based on differences in hourly wages between individuals of the same age with a college and a high school diploma.[4] This poses advantages and disadvantages. On the one hand, an advantage of this data is that it compares individuals who were subject to the same influences on their decision to attend post-secondary education. On the other hand, a potential disadvantage is that it ignores any possible difference in labour experience between people of the same age but with different educational attainment. To overcome this possible threat is that, in the empirical specification, I account for systematic age effects in the wage gap.

To assess the effect of city size on the college wage gap across the age profile I need a time-consistent definition of city. In this project I use the 2013 Core Based Statistical Area (CBSA) definition of city. The CBSA definition involves taking one or more counties (or equivalents) that are attached to an urban center, including all the adjacent counties socioeconomically tied to that urban center by commuting routes. Using this definition yields 917 cities in the United States. I use the CBSA definition instead of other alternative definitions of cities, such as Commuting Zones, because it better suits the comparisons across cities.

One challenge associated with the use of census data is that its geographic units rarely line up with the CBSA definition. To assign individuals in each decennial census to a CBSA I create a crosswalk between Public Use Microdata Area (PUMA) and CBSAs. If a PUMA spans over multiple CBSAs, I create population allocation factors. Thereafter, I interact these factors with the census weight. This means that some individuals are counted multiple times in the data, but with overall weights that still add to their contribution to the US population.

[3] For those years where the minimum wage changes during the year I use the minimum wage that was in place in January 1st.

[4] I stacked individuals in groups of 5 years. For example, I treat all workers that are between 36 and 40 as if they have the same age.

3 Preliminary Data Patterns

The evolution of local labour markets over time has been different for young workers compared to old workers. In this Section, I start by noting that the college wage premia and the elasticity of wages with respect to city population increase differently over time for workers at different stages of the work-cycle. Then, I turn my attention to the geographical variation of the college wage gap and the relative college labour supply. More populous cities have traditionally had larger college wage premia and a more college intensive labour supply. Finally, I present evidence on how the decadal change in the college wage gap evolved faster for more populous cities.

Table 1 broadly outlines the motivation for this paper. It shows that both the college wage premium and the elasticity of hourly wages with respect to city size have evolved differently over time for workers in different stages of the work cycle. The first three columns show the average college wage premia for every decade since 1980. The last three columns show the average wage elasticity with respect to the 1980's city population. The estimates use log hourly wages of full-time male workers who work more than 50 weeks in a year as the dependent variable. All cells in the table correspond to a different regression and are weighted by the census weight interacted with the labour supply.[5]

The college wage premia has increased in every decade since 1980. For workers across all age groups the level of the college wage premia rose by 0.26 between 1980 and 2019. The table also shows that the college wage premia rises over the work-cycle, as seen by Card and Lemieux (2001).[6] However, the difference in the level of the college wage premia across the work-cycle has shrank over time. The convergence in the college wage premium across the age profile can be explained by workers in early stages of their career who faced larger increases in the college wage premia. For example, between 1980 and 2019 the wage premia for young workers more than tripled whereas for old workers it increased by 30%.

Initially dispersed across the work-cycle, the wage elasticity with respect to city size

[5] I define labour supply as the interaction between usual hours worked in a week and the total number of weeks in a year (52). See Section 2 for more details.

[6] Despite the different data sources (Census and ACS vs. CPS) and definition of wages used (hourly vs weekly wages), the estimates of the college wage premia in Table 1 follows closely those in Table 1 in Card and Lemieux (2001). For example, I estimate that for workers between 26 and 30 in 1980 the college premia was 0.139 whereas Card and Lemieux estimate a premia of 0.111. For this same age group in 1990 I estimate a premia of 0.307 while Card and Lemieux's estimates is 0.331.

has converged across different age groups over time. The results for overall workers are in line with the short time estimates of Duranton and Puga (2019) and indicate that the city size wage premium increased from 1980 to 1990, but remained relatively stable at 5% after the 2000s. The elasticity for older workers follows closely the evolution of the elasticity for all workers. For younger workers the elasticity has converged to the overall average: after the rapid increase between 1980 and 1990, the elasticity remained stable at 4.7%. This hints that in recent decades larger cities have become relatively more attractive for younger workers than in 1980.

The manner in which the college premium interacts with the city size wage premium is a field economists have paid less attention to. Figure 1 highlights the positive relationship that exists between the college-to-non-college wage gap and city size, regardless of the worker's stage in the work-cycle. This relationship is referred to as the college bias of agglomeration externalities. The figure is constructed using average hourly wages for college and non-college graduates in each of the 917 cities used in this paper. Each panel is the predicted value of a kernel-weighted local polynomial smoothing of the college wage gap on the (log) 1980 CBSA population.

Figure 1 reveals two important insights. Panel (a) shows that, while larger cities have traditionally had a larger college wage gap, the level and slope of this city-size-college-wage-gap relationship rose consistently in each decade. This pattern prevails in each stage of the lifecycle, as shown in Panels (b) and (c). In 1980, for example, the overall college wage gap in the most populated city was about 40% larger that the average college wage gap in the least populated city. By 2019 this gap had increased to 75%.

Perhaps most strikingly, Figure 1 depicts that the change in the level and slope of the city-size-college-wage-gap relationship increased faster for younger workers compared to older ones. For example, in 1980 the average college wage gap in the most populated city was roughly the same as the college wage gap in the city with the median population regardless the age of the worker. By 2019 the ratio was more than 50% larger in the largest city for young workers and less than 20% for old workers. These differing patterns suggest that the college bias of agglomeration externalities evolved differently depending on the age profile.[7]

[7] The most populated city is New York City. Half of the U.S. population lives in a city with a population smaller than Orlando City.

The aforementioned patterns are not caused by a reduction in the relative supply of college workers over time. Indeed, in the last fifty years the relative supply of college workers has increased. Figure 2 shows that the level and slope of the relationship between the relative college labour supply and city population has steadily increased with each passing decade.[8] The three panels of Figure 2 show that more populous cities have a larger relative labour supply of college workers. Panel (a) illustrates the urban gradient in the relative labour supply of college workers despite their age and how it has increased over time. This pattern is in line with previous studies such as Autor (2019), Costa and Kahn (2000), Diamond (2016), Florida (2002), Glaeser and Mare (2001), and Moretti (2013). Panel (b) shows that for young workers the gradient of this relationship has steepened with time, specially in the last decade. Surprisingly, Panel (c) shows that the relative supply of college workers in late stages of their careers peaked in 2010 and decreased afterwards to roughly the same level as in 2000. These patterns suggests that college graduates locate disproportionately in larger cities compared to high school graduates.

Table 2 quantifies the patterns shown in Figures 1 and 2. In other words, it evaluates the changes in the relationship between the college wage gap or the relative college labour supply and city size over time. The first three columns show the elasticity of the college wage gap with respect to the 1980 city size, whereas the last three columns show the elasticity between the relative labour supply of college workers and the 1980 city size. All entries are weighted by the 1980 city population.

Table 2 shows that the elasticity of the college wage gap for all workers with respect to city population has increased over time. After an initial decrease, the elasticity rises steadily in each decade since 2000. Between 1980 and 2019 the level of the college wage gap elasticity with respect to city size increased on average by 2.7 percentage points.

Historically more college intensive, more populous cities have also employ more college workers with each passing decade. This result is referred to as the *great divergence* (Moretti, 2012). Columns 4-6 of Table 2 show that the elasticity of the relative college supply has increased over time. Larger cities consistently employ more skilled workers and have become increasingly more college intensive with each passing decade.

More populous cities have also seen a larger decadal increase in the college wage gap

[8] The relative college labour supply of a city makes reference to the total hours worked in a year by college graduates divided by the the total hours worked by high school graduates in a year.

and in their relative college labour supply. Table 3 presents regressions of decadal changes in the college wage gap or relative labour supply on city size and decadal dummies. Mirroring Baum-Snow et al. (2018), these results are intended to capture the average decadal change in the elasticities of the mentioned dependent variables with respect to city size. The change in the elasticity of the college wage gap with respect to city size significantly increased by 0.07 over each decade, regardless of the worker's age.

Furthermore, Table 3 shows that more populated cities have larger growth in the elasticity of the relative labour supply of college graduates. Thus, this change is not homogeneous across age groups. Over each decade, the elasticity of the relative skilled labour supply for young workers increased by 0.042. For old workers the same effect was about 0.02. This pattern hints that young college graduates are disproportionately relocating to larger cities.

The patterns presented in this section make a critical point: the increase of the urban bias of agglomeration externalities over time is stronger for younger workers. In order to interpret the pronounced increase of the college wage premia in the United States it is necessary to understand why the rise in the agglomeration externalities across cities differs across the workers' work-cycle. The primary hypothesis advanced in this paper is that over the last fifty years, the faster increase in the college wage gap experienced by cities with larger population is due to a shift in the occupational structure of local labour markets.

4 Theoretical Framework

In this section I develop the motivating model used in this paper and its relationship to the estimating equation.

4.1 The Model

Patterns in the data discussed in Section 3 are consistent with the idea that agglomeration externalities benefit college graduates, that this effect is not homogenous across workers of different ages, that these effects have strengthen with time, and that bigger cities have become relatively more skilled over each passing decade. To formalize these patterns I develop a model of city aggregate production with age-group specific labour supplies. I build on Card and Lemieux's (2001) and Baum-Snow et al.'s (2018) models of nested constant elasticity of substitution production functions. The model relaxes the hypothesis of perfect substitution and homogenous impact of agglomeration externalities across age groups. A way of relaxing these assumptions is to assume that aggregate output depends on two CES subaggregates of high-school and college labour that are impacted differently by agglomeration externalities:

$$S_{ct} = \left[\sum_j D_c^{\mu_{tj}^s} S_{ctj}^{\eta} \right]^{\frac{1}{\eta}} \tag{1}$$

and

$$U_{ct} = \left[\sum_j D_c^{\mu_{tj}^u} U_{ctj}^{\eta} \right]^{\frac{1}{\eta}} \tag{2}$$

where η is a function of the partial elasticity of substitution σ_A between age groups j with the same level of education ($\eta = 1 - \frac{1}{\sigma_A}$).[9] D_c is a measure of city population density. The parameter μ_{jt}^g reflects the agglomeration forces of educational group $g \in \{s, u\}$. For every age group j, college biased agglomeration externalities requires that $\mu_{jt}^s > \mu_{jt}^u$. Given that the college wage gap increases with city size, I expect this condition holds in the data.

City c's aggregate output function at time t is a function of college and non-college workers, and the technological efficiency parameters A_{ct}, θ_t^s, and θ_t^u:

$$Y_{ct} = A_{ct} \left(\theta_t^s S_{ct}^{\rho} + \theta_t^u U_{ct}^{\rho} \right)^{\frac{1}{\rho}} \tag{3}$$

where ρ is a function of the elasticity of substitution σ_E between the two educational groups ($\rho = 1 - \frac{1}{\sigma_E}$). In this setting, the marginal product of labour for a given age-education group depends on a national skill level shock, the agglomeration externalities

[9] Card and Lemieux (2001) also have time invariant relative efficiency parameters in the CES sub-aggregates of high-school and college labour (Equations 1 and 2 of the paper). Since these parameters don't change with time they will not appear in the equation to estimate. Therefore, I don't add them in the model.

it faces in city c, the groups own labour supply in city c, and the aggregate supply of labour supply in its education category in city c. Then, the marginal product of a college worker in city c and age group j is:

$$\begin{aligned} \frac{\partial Y_{ct}}{\partial S_{ctj}} &= \theta_t^s \Psi_{ct} S_{ct}^{\rho-1} S_{ct}^{1-\eta} D_c^{\mu_{tj}^s} S_{ctj}^{\eta-1} \\ &= \theta_t^s \Psi_{ct} S_{ct}^{\rho-\eta} D_c^{\mu_{tj}^s} S_{ctj}^{\eta-1} \end{aligned} \tag{4}$$

where

$$\Psi_{ct} = A_{ct} \left(\theta_t^s S_{ct}^{\rho} + \theta_t^u U_{ct}^{\rho}\right)^{\frac{1}{\rho}-1}$$

Similarly, the marginal product of a non-college worker in city c and age group j is

$$\frac{\partial Y_{ct}}{\partial U_{ctj}} = \theta_t^u \Psi_{ct} U_{ct}^{\rho-\eta} D_c^{\mu_{tj}^u} U_{ctj}^{\eta-1} \tag{5}$$

I will assume that labour markets are competitive.[10] Therefore, for each age group j the relative marginal product of workers of different skill groups equates the relative wages. Hence, Equations 4 and 5 imply:

$$\begin{aligned} \underbrace{log\left(\frac{w_{ctj}^s}{w_{ctj}^u}\right)}_{log(r_{ctj})} = log\left(\frac{\theta_t^s}{\theta_t^u}\right) &+ (\rho-\eta) log\left(\frac{S_{ct}}{U_{ct}}\right) \\ &+ (\mu_{tj}^s - \mu_{tj}^u) log\left(D_c\right) + (\eta - 1) log\left(\frac{S_{ctj}}{U_{ctj}}\right) \\ = log\left(\frac{\theta_t^s}{\theta_t^u}\right) &+ \left(\frac{1}{\sigma_A} - \frac{1}{\sigma_E}\right) log\left(\frac{S_{ct}}{U_{ct}}\right) \\ &+ (\mu_{tj}^s - \mu_{tj}^u) log\left(D_c\right) - \left(\frac{1}{\sigma_A}\right) log\left(\frac{S_{ctj}}{U_{ctj}}\right) \end{aligned} \tag{6}$$

Since we are interested in the heterogeneous evolution of agglomeration externalities across different age groups in city c I will totally differentiate Equation 6. This gives us

[10] This assumption may not be innocuous. There is evidence that larger cities have more competitive labour markets (Azar et al., 2020). However, exploring the monopsony power of firms across cities is out of the scope of this project.

the central equation of the paper:[11]

$$\mathrm{d}log(r_{cj}) = \mathrm{d}log\left(\frac{\theta^s}{\theta^u}\right) - \frac{1}{\sigma_E}\mathrm{d}log\left(\frac{S_c}{U_c}\right) + (\mathrm{d}\mu_j^s - \mathrm{d}\mu_j^u)log\left(D_c\right) + \left(\frac{1}{\sigma_A}\right)\left(\mathrm{d}log\left(\frac{S_c}{U_c}\right) - \mathrm{d}log\left(\frac{S_{cj}}{U_{cj}}\right)\right) \tag{7}$$

The derivation can be found in the Appendix Section C. Equation 7 incorporates all the mechanisms that may drive a city's change in wage inequality. In city c, the city wage gap increases if (i) there is an increase in the relative technological efficiency parameters that favours college workers, (ii) the overall relative supply of college workers in a city decreases over time, (iii) the college bias of agglomeration externalities increases over time, and (iv) workers of different age groups are imperfect substitutes (i.e. $\sigma_A < \infty$) and the increase of the overall relative supply of college workers is higher than the increase in the relative supply of skilled workers in age group j (i.e., group j is relatively scarce compared to the city's level).

Figure 2 shows that larger cities have seen an increase in the relative supply of skilled workers and have also seen an increase of the college wage gap over time. Hence, Equation 7 suggests that college biased agglomeration externalities are a relevant factor to explain the increase in the college wage gap over time.

4.2 Matching the Model to the Data.

The primary interest of this paper is to estimate the evolution of the college bias of agglomeration externalities across age groups. In the model these parameters are captured by $(\mathrm{d}\mu_j^s - \mathrm{d}\mu_j^u)log(D_c)$. These parameters change over time, age group, and city density. I parametrize these effects as an interaction between a year dummy, an age group dummy, and city c's 1980 population.[12] The parameter $\mathrm{d}log\left(\frac{\theta^s}{\theta^u}\right)$ varies across decade and it is parametrized by a time fixed effect.[13] The variables $\mathrm{d}log\left(\frac{S_c}{U_c}\right)$ and $\mathrm{d}log\left(\frac{S_{cj}}{U_{cj}}\right)$ come from the data and correspond to the decennial change in the (log) ratio of total hours worked

[11] I will drop the t subscript. I will reintroduce them in the estimating equations.

[12] Figure 1 suggests that the effect agglomeration externalities have on the college wage gap is approximately linear. Consequently, all specifications use a linear measure of population density. The main measure is (log) 1980 city population. As a robustness check I also use 1980 city density (population/area) as an alternative measure. The results are robust to this alternative measure.

[13] Table 5, column 2 shows that the results are robust to the use of linear time trends as alternative parametrization.

in city c by college and non-college workers, and the ratio of total hours worked in city c by college and non-college workers in age group j, respectively. I write the empirical version of Equation 7 as:

$$\Delta log(r_{ctj}) = \sum_{h\in J}\sum_{\tau\in T}\underbrace{\mu_{h\tau}Pop_{c\tau h}}_{(\mathrm{d}\mu_j^s - \mathrm{d}\mu_j^u)} + \underbrace{\delta}_{\frac{1}{\sigma_A}-\frac{1}{\sigma_E}}\ \Delta log\left(\frac{S_{ct}}{U_{ct}}\right) \tag{8}$$
$$+ \underbrace{\gamma}_{-\frac{1}{\sigma_A}}\ \Delta log\left(\frac{S_{ctj}}{U_{ctj}}\right) + \underbrace{\lambda_t}_{\mathrm{d}log\left(\frac{\theta^s}{\theta^u}\right)} + \psi_j + \epsilon_{ctj}$$

where $\Delta log(r_{ctj})$ correspond to the decadal change in the log college wage gap for age group j in city c and period t, and $Pop_{c\tau h} = \mathbb{1}(\tau = t) \times \mathbb{1}(h = j) \times log(D_c)$ captures the heterogeneous impact agglomeration externalities have across age group and time. J and T are the sets of ages and years used, respectively. λ_t and ψ_j are year and age group fixed effect, and ϵ_{ctj} is an approximation error. I include age group fixed effect to control for systematic differences between college and high school workers of age j, such as work experience. I include under each term its counterpart in Equation 7. The objects of interest are the coefficients $\mu_{h\tau}$s. It is easy to relate the estimates δ and γ to the elasticity of substitution between college and non-college workers, and between different age groups. Specifically, the model predicts that $\delta = \left(\frac{1}{\sigma_A} - \frac{1}{\sigma_E}\right)$ and $\gamma = -\left(\frac{1}{\sigma_A}\right)$.

Taking Equation 8 directly to the data would result in biased estimates because the relative supply of college and non-college labour is an equilibrium outcome. In other words, $\Delta log\left(\frac{S_{ct}}{U_{ct}}\right)$, $\Delta log\left(\frac{S_{ctj}}{U_{ctj}}\right)$, and $\Delta log(r_{ctj})$ are simultaneously determined making a credible identification difficult. Therefore, identification of parameters requires exogenous variation in changes in the relative supply of skill across cities.[14] I will use past immigrants' enclaves and contemporaneous immigration shocks as source of exogenous variation. The instruments are explained in the following section.

[14] Another source of endogeneity would be a positive labour demand shock that attracts college workers into a city disproportionately relative to non-college workers. A positive shock to college workers will increase the unobserved component as well as the relative supply of college workers.

4.3 Exogenous Variation in Local Labour Supply

I will follow Lewis (2011) and Baum-Snow et al. (2018) and use the location where immigrants settled in 1970 to predict the immigrant's relative labour supply for each CBSA between 1980-2019. The prediction will be used as an instrument of the relative supply of college and high-school graduates in each city. The idea behind the instrument is that for reasons not related to the contemporaneous conditions of local labour market, immigrants are more likely to settle in locations with a relatively larger population of their country of origin. The change in the relative supply of immigrants impacts city c's labour supply without affecting the demand for labour.[15]

The predicted quantity of college workers in city c and time t follows:

$$\begin{aligned} \hat{S}_{ct} &= \sum_o \frac{M^o_{c,1970}}{M^o_{1970}} \times S^o_t \\ &= \sum_o s^o_{c,1970} \times S^o_t \end{aligned}$$

where $\frac{M^o_{c,1970}}{M^o_{1970}}$ is the share of the immigrant stock from country o living in city c in 1970 regardless of skill level. This captures the idea of immigrant enclaves. Notice that the shares are not skill specific. Also, the shares do not sum to 1 within a city. S^o_t is the number of immigrants from country o with a college degree who settled anywhere in the U.S. during the period of analysis. The immigrants' inflows are skill specific.

The predicted quantity for non-college workers is analogous and follows:

$$\hat{U}_{ct} = \sum_o s^o_{c,1970} \times U^o_t$$

Similarly, the predicted quantities for workers in age group j are:

$$\hat{S}_{ctj} = \sum_o s^o_{c,1970} \times S^o_{tj}$$

[15] An implicit assumption is that immigrants and native workers are not complements. If so, immigrants arrival would increase the productivity of native workers, thus increasing the overall demand.

and

$$\hat{U}_{ctj} \quad = \quad \sum_o s^o_{c,1970} \times U^o_{tj}$$

where S^o_{tj} and U^o_{tj} are the number of new immigrants from country o and age group j that arrived anywhere in the U.S. in period t.

The main equation of this paper looks at changes in the college wage gap over time. To incorporate this into my instrument I take first differences in the predicted quantities of immigrants relative labour supply. In particular, the predicted instrument follows:

$$\Delta_t ln(\widehat{\frac{S_{ctj}}{U_{ctj}}}) = ln(\widehat{\frac{S_{ctj}}{U_{ctj}}}) - ln(\widehat{\frac{S_{ct-1j}}{U_{ct-1j}}})$$

I check the predictive power of the instrument by estimating the relationship between the growth in the predicted immigrant college labour supply in each city and the growth in the college labour supply. In particular, I estimate:

$$\Delta_t ln\left(\frac{S_{ctj}}{U_{ctj}}\right) = \nu_t + \varpi_j + \alpha_1 \Delta_t ln\left(\widehat{\frac{S_{ctj}}{U_{ctj}}}\right) + \alpha_2 ln\left(\frac{S^{imm}_{ct-1j}}{U^{imm}_{ct-1j}}\right) + \alpha_3 ln(D_c) + u_{ctj} \tag{9}$$

where Δ_t denotes the difference between periods t and $t-1$, and the superscript imm indicates the immigrants' relative labour supply in the last period. When looking at the overall relative labour supply I drop the j subscript and the age group fixed effect ϖ_j. Following Lewis (2011) and Baum-Snow et al. (2018), I include lagged relative supply of college immigrants to remove any potential correlation between period $t-1$ relative immigrant labour supply and changes in the contemporaneous college labour supply. This also controls for Jaeger et al.'s (2018) critique that local labour markets take time to adjust to long term labour supply shocks.

Table 4 shows the results of Equation 9. It suggests that the predicted relative immigrant labour supply is a strong predictor of the relative labour supply. These results indicate most of the identifying variation comes from labour supply shocks to each age group. Each column of the table corresponds to a different regression of the change in the (log) relative supply of college workers in the predicted number of immigrant college labour supply. All the regressions are weighted by the 1980 CBSA population and clustered by CBSA. A 10 percent increase in the predicted immigrant college labour supply leads to an estimated 7 percent increase in relative skilled workers for an specific age group, and to a 0.97 percent increase in the overall relative college labour supply.

The identification assumption of the instruments is that contemporaneous shocks to productivity experienced by each CBSA are independent of the shocks in 1970 that drew immigrants to each CBSA, conditional on city size, age, and year. This assumption is equivalent to Goldsmith-Pinkham et al.'s (2020) interpretation of the exogeneity of the shares. The instruments would not meet the assumption if, for example, college educated immigrants in 1970 chose in which CBSA to settle because they anticipate a change in the growth of the college wage gap.

One of the recent critiques of immigrants enclaves' instruments is that it conflates the short and long term effect of the immigrants shock on the relative supply of workers (Jaeger et al., 2018). If local labour markets takes time to adjust to demand shocks, then the error term in equation 8 would also include lagged shocks that reflect the ongoing general equilibrium adjustment of past immigration supply shocks. Although I cannot rule out this as a possible concern, three reasons make me think this is not a threat to the paper's results. First, the use I give to this instrument differs from the traditional use. I use it because immigrant's migrations impacts the total relative labour supply in a city, whereas traditionally it has been used to estimate the change in the specific immigrant relative labour supply. In other words, I use it to instrument the overall relative labour supply for each age group instead of the relative labour supply of immigrants. Second, the shares come from ten years before my sample. Jaeger et al. (2018) argue that with low frequency data, as it is the decennial census, the adjustment bias is smaller. The authors implicitly supports the idea that after two decades the adjustment bias is negligible.[16] Hence, the adjustment bias created by the use of 1970's shares should be negligible. Finally, I include the relative labour supply at the beginning of each period in order to address the concerns this paper points out. This inclusion doesn't change the point estimates nor the fit of the model.

A possible concern with the instrument is that natives and immigrants can move

[16] They do this by only using one lagged variable while using decennial census.

between cities in response to new immigrant influx (Borjas, 2003). Selective out-migration of incumbent workers can undone the increase in the labour supply of the immigrants influx, weakening the predictive power of the instrument. However, Card (2001) and Card and DiNardo (2000) find no evidence of native out-migration as a response of immigrant inflow. In any case, the out-migration of incumbent workers would attenuate the first stage but would not be a problem for the identification. Table 4 shows that the instrument is a strong predictor of the changes in the relative supply of college workers.

A second possible issue with the use of the predicted immigrant stock is that immigrant labour force couldn't be a perfect substitute of native labour force as it is shown in Card (2009); Ottaviano and Peri (2012); Dustmann et al. (2012), and Manacorda et al. (2006). This concern doesn't pose a threat to identification unless the labour supply of native college workers is unrelated to immigrant inflows and vice versa. In other words, if college immigrants were a better substitute of non-college natives, and vice versa. However, Borjas et al. (2008) argue that some of these lack of substitutability hinge on the way the sample of working individuals is constructed. Since the growth in the predicted immigrant labour supply is a strong predictor of the growth of the labour supply, I'm not too concerned that the different degrees of substitutability have much influence on the results.

5 Results

The objective of this paper is to document the heterogeneous effect city population has on the college wage gap across the work-cycle. This section presents the estimates of agglomeration externalities in the college wage gap for workers of different age. I find that more populous places have a larger college wage gap, but this effect is not distributed homogeneously across the work-cycle. Young workers have face a larger increase in the college bias of agglomeration externalities over time. In addition, the specification used also allows me to recover the elasticity of substitution between workers of different age groups and between workers of different educational attainment. I estimate an elasticity of substitution between workers of different age groups in line with the labour literature. The estimated elasticity of substitution between college and high school workers, however, is larger than the usual estimates in the literature.

Table 5 report the IV estimates of the elasticity of substitution between college and non-college workers and the elasticity of substitution between different age groups. Column 1 includes year and age group fixed effect but it doesn't control for the lagged relative supply of college immigrants discussed in Section 4.3, whereas Columns 2-4 include this variable. Column 2 is estimated using a linear time trend instead of time fixed effect in addition to an age group fixed effect. Column 3 is estimated using only a time fixed effect. Column 4, the preferred specification, includes time and age group fixed effect. For every specification the first stage F-statistic is well above 10, the usual rule of thumb. The first stage of the preferred specification can be found in Table A.3.

I find that an increase in the relative supply of college workers decreases the college wage gap – as expected in a supply and demand model. The results in Table 5 suggest that an exogenous increase in the relative supply of college workers of age group j leads to a statistically significant decrease in the age-group college wage gap in all the specifications. The decreases range between -0.219 and -0.233. Holding the agglomeration externalities and composition constant, a 10% decadal increase in the relative supply of college workers in age group j is associated with an average decadal decrease in the college wage gap of 2.2%.

Results for the the overall relative supply of college workers are not statistically different from zero. All of them are negative, suggesting that an increase in the relative supply of college workers impacts all age groups. However, under none specification the estimate is statistically significant. This favours the hypothesis that workers are not perfect substitutes across the age profile. Once accounted for the age-group specific relative supply of college workers, the impact workers of different age have on age group j's wage gap is statistically indistinguishable from zero. Also, the inclusion of the lagged relative supply of college immigrants doesn't change the point estimates of the relative supply of college workers, and the estimate of this variable is not statistically different from zero in any specification.

I can recover the elasticity of substitution between different age groups (σ_A), and between college and non-college workers (σ_E) from Equation 8. The estimates imply an elasticity of substitution between different age groups in the range of 4.2 and 4.6. These results are in line with the results in Card and Lemieux (2001), who find that the elasticity of substitution across age groups ranges from 4 to 6. This result back up the assumption of imperfect substitution between workers of different age groups (i.e. $\sigma_A = \infty$).

The results in Table 5 imply an elasticity of substitution between college and non-college workers of 3.65. The 95% confidence interval locates this coefficient between 1.7 and 5.6. The labour evidence suggests that this elasticity ranges from 1.2 to 2 (Katz and Murphy, 1992; Card and Lemieux, 2001; Bound and Johnson, 1992; Ciccone and Peri, 2005). The estimates I find double the results of previous studies. However, most of the existing literature uses data prior to 2000. Baum-Snow et al. (2018) find that the elasticity of substitution rises with each passing decade.[17] Consequently, the larger estimates I find are the result of using two extra decades to the usual years used in the labour literature.

Table 6 presents the main results of the paper. The entries in Table 6 provide a variety of information about the evolution of the college bias of agglomeration externalities. Comparisons down a column of the table show the changes over time in college bias agglomeration externalities for specific age groups. Comparisons across rows reveal the age profile of the college bias agglomeration externalities in each decade. These estimates come from the specification used in Table 5, column 4. The estimates for the other specifications can be found in Tables A.4, A.5, and A.6. The results for the alternative specifications are very similar to the results in Table 6.

Consistent with Figure 1, I estimate that the college bias of agglomeration externalities is positive for all age groups and decades. The results are statistically different from zero for almost every age group-decade. The only non-significant estimates are those for 45-50, 51-55, and 56-60 years old workers in 1990. Holding the relative supply and composition constant, the increase in the wage gap for workers between 26 and 30 located in the city with the largest population would be 3.7% and 7.1% larger than the median city in 1980 and 2019, respectively. For workers between 56 and 60 this same effect would be 2.4% and 3.4% larger.

One advantage of the econometric design used is that I can empirically test the as-

[17] I also replicated the results in Table 5, column 4 of Baum-Snow et al. (2018). Although I'm using a different definition of city, a different sample of workers (I use full-time full-year male workers whereas they use all male workers), and an extra decade, my estimate of the elasticity of substitution between college and non-college workers is similar to theirs (3.3 vs 2.3). In addition, the estimates of the other variables have the same sign and magnitude.

sumptions made in the theoretical model. Specifically, I can test if the college bias of agglomeration externalities differs across the age profile and if this effect changes with time. The last column in Table 6 shows the p-values of the joint test that in a given decade all age groups' coefficients are the same. In each decade we can reject the null hypothesis of constant effect across the age profile. This hints that the benefits of cities are not shared equally across different age groups. Interestingly, the largest increase in the benefit of agglomeration externalities is in workers younger than 40 years old. This results could explain the urbanization of younger cohorts that the U.S has seen in the last decades.

I also test the assumption that agglomeration externalities have changed over time. The last row in Table 6 shows the p-values of the joint test that the effect for each age group is constant between decades. For each age group we can reject the null hypothesis that the effect is constant across decades. However, for the youngest and oldest workers the statistical significance of the test is 5%. For the other age groups the results are different at the 1% level.

6 Robustness

In this section I perform a battery of robustness exercises to the results in Section 5.

6.1 Sorting

One key concern in this paper is the issue of sorting. In particular, changes in unobserved characteristics of workers in a city over decades may be correlated with the population of that place. For example, if in each passing decade there is an increase in the amount of college workers with high unobserved abilities that move to larger cities. In this section I present evidence that sorting is not driving the results of the paper. There is no direct test for sorting on unobserved characteristics without panel data. However, I provide two indirect tests for this concern.

First, in Table A.7 I re-estimate Table 6 using the subsample of workers who are located in their birth state. This test rules out the possibility of sorting between states. The results support the evidence that the college bias of agglomeration externalities is larger for workers in early stages of their career and that it increases over time. The estimates for 1990 are not statistically different from zero. Noteworthily, the statistically significant estimates are smaller in magnitude for all ages and years. This hints that the sorting of college graduates partly explain the increase in the college bias of agglomeration externalities across the age profile. Interestingly, for workers over 50 years old in 2019 the coefficients of this sample lie within one standard deviation from the original estimates, suggesting that sorting is stronger in younger cohorts. However, the overall take away of Section 5 remains unchanged: college bias of agglomeration externalities has a different impacts on workers across the work-cycle, and this effect has increased over each decade.

The second indirect test is to see how the predicted error term of a mincer equation varies with city population. If sorting is true, then larger cities should have workers with a higher unobserved component of the wage. Figure B.2 plots the binscatter plot of the error of a regression of (log) hourly wage on educational attainment, age, city, decade, country of origin, race, and industry dummies. Notice that the level of the predicted error is very small. Scaled by a factor of 100,000,000 the plot shows that the predicted unobserved component increases with population. I don't consider this a threat to results of the paper because of the scale of the results. Although there is a clear positive relationship, the scale needs to be multiplied by 100 million in order to see the pattern arise. Furthermore, even after the re-scale the level of the relationship is very close to zero.

6.2 Calibrating the Elasticities of Substitution.

The main purpose of this paper is to estimate the change in the college bias of agglomeration externalities across the age profile over time. These estimates comes from a regression that includes the ratio of college to non-college workers which are simultaneously determined with the college wage gap. To overcome this threat, in Section 5 I use the exogenous variation in labour supply that arises from immigrants' enclaves. An alternative approach to deal with the endogeneity is to calibrate σ_A and σ_E to match values previously found in the labour literature. In this section I will use the later approach. To check if the results are sensible to calibrated parameters, I run 4 different specifications. I calibrate the elasticity of substitution between different age groups (σ_A) to equal 4.5

and 5.5. The elasticity of substitution between workers with and without a college degree (σ_E) is calibrated to 1.4 and 1.7. Both calibrated elasticities correspond to values the literature has found to be in the middle of the range.[18]

Tables A.8, A.9, A.10, and A.11 present the results for the constrained estimation. The main difference is that the results are between 1.5 and 3 times larger than those in Section 5. However, these tables support the main take away of the paper: city size impacts the college wage gap differently across the age profile. Workers under 40 years old present a larger city size gradient. Moreover, the effect increase over time. Overall, I interpret the evidence in this Section as broadly supporting the notion that the college bias of agglomeration externalities impacts differently the college wage gap across age groups and that it changes over time.

6.3 Efficiency Units.

The results in Section 5 cannot disentangle the effect of changes in the price of observed characteristics from changes in the composition of the workforce within college and non-college workers. Although both effects are associated with agglomeration externalities, unraveling the weight of each force sheds lights on the nature of the change in city size gradient of the college wage gap. Thus, I construct efficiency units to control for the composition effect.

To construct the number of efficiency units each worker contributes I follow Baum-Snow et al. (2018). I separately regress the log hourly wage in 1980 on age, race, industry, country of birth, and CBSA dummies for workers with and without a college degree. I interpret the coefficients of these two regressions as the productivity-price of each observable characteristic in 1980. I then use the coefficients of these two regressions to predict the wage of each worker for the decades after 1980. To predict the number of labour efficiency units associated with each worker if they worked in a reference CBSA I interact the hours worked by each worker by its predicted (exp) hourly wage. I adjust the new efficiency units labour supply to match the first moment of the original labour supply distribution. I use the coefficients of the 1980 sample for later years to avoid the issue of endogenous changes in prices and composition. Finally, I aggregate the new microdata

[18] I took the range of σ_A from the estimates in Card and Lemieux (2001). The value of σ_E comes from Card and Lemieux (2001); Katz and Murphy (1992), and Ciccone and Peri (2005).

to match the data used in the paper.

Table A.12 presents the estimates using efficiency units. As expected, the coefficients are smaller in magnitude. Interestingly, the coefficients for the year 1990 are not statistically different from zero. This suggests that the results in Table 6 arise from changes in the composition of the workforce in each city. For the rest of the years in the sample the estimates are positive and increasing across decades. As in the previous section, the coefficients are stable up to the 41-45 years age group and then decreases with age. We can also reject the null that neither the effect is homogeneous across years nor across the age profile. Overall, the evidence in this subsection supports the notion that the college bias of agglomeration externalities impacts the college wage gap across age group and it changes over time after controlling for compositional shifts.

6.4 Big Cities.

A possible concern in Table 5 and 6 is that the estimates of the average college wage gap for each age group in a city is driven by outliers. To estimate the college wage gap for each city in every decade I need a large number of workers in each city, age group, and educational group. This is a very demanding condition for the data, specially for small cities. To indirectly control for this concern I weight the average college wage gap by the 1980 CBSA population. Therefore, those potentially biased estimates of the wage gap 'weight less' in the total estimation. To directly test if the results of the paper are potentially biased by those observations, I run Equation 8 using cities with a relatively large population. More specific, Table A.13 presents the estimates using the subsample of cities that had a 1980 population larger than 25,000, 50,000, 75,000, and 100,000. The results resemble those in Table 5, with slightly lower estimates of σ_A and σ_E. For example, the estimates for the elasticity of substitution between different age groups ranges from 4.35 to 4.6.

The estimates for the change in the college bias of agglomeration externalities are presented in Tables A.14 - A.17. For each one of the four specifications the results are virtually unaltered: all the estimates are slightly higher but lie within one standard deviation from those in Table 6. In each specification I can reject the hypothesis that the effect city size has on the college wage gap is homogeneous across age groups. However, for those workers who were between 26-30, 31-35, and 51-55 in cities with a population bigger than 75,000, I cannot reject the null hypothesis that the effect varies across decades. These estimates imply that the estimates found in Section 5 are not driven by outliers in the data.

7 Occupational Shift

In this section I show that the divergent shift from highly routinary tasks is the cause of the heterogeneous evolution of the college bias of agglomeration externalities across the work-cycle. Firstly, I show that the urban gradient of wages and its evolution over time differs amid educational groups. Figure 3 shows that the level and slope of the relationship between average wage and city population for college graduates has increased with each passing decade. College workers in all age groups have faced roughly the same evolution of the wages-city-population relationship. Instead, for high school the same relationship has decreased in level and slope over time. These patterns suggest that the increase in the urban gradient and level of the college wage gap, shown in Figure 1, is due to both the improvement for college graduates and the deterioration for high school graduates' urban wages.

Figure 3 reveals that the urban gradient of high school graduates' wages at different stages of the work-cycle diverged over time. For workers with just a high school diploma that were between 26-30 in 1980 there is a clear positive relationship between city population and wages. In 2019, not only the average wage in all cities was considerably lower, but also the relationship between city population and wages was flat. Conversely, in each decade since 1980 workers in later stages of their careers have a clear positive relationship between wages and city population. They also faced a smaller decrease in the wage level with each passing decade. This suggests that agglomeration externalities are also experience biased. However, studying the interaction of experience and city population is outside the scope of this paper.

Secondly, I show that the aforementioned differences in the wage evolution are explained by a shift in the structure of the tasks performed by workers at different stages of their work-cycle. Figures 4 and 5 illustrate the employment share of each educational-age group in occupations with high and low routinary task index (RTI) scores. I use the RTI definition constructed by Autor and Dorn (2013) because it maps each occupation's manual, abstract, and routine inputs into one comparable index.[19] An occupation is catalogued as highly routinary if its RTI score belongs to the top tercile of the RTI distribution. Conversely, low routinary occupations are those whose RTI score belongs to the bottom tercile of the RTI distribution.

Additionally, Figures 4 and 5 reveal three new insights. First, they show that employment shares are not population neutral. The U.S. labour force has traditionally been more intensive in low RTI jobs but the city population gradient of this share is negative. However, for college workers this relationship flattens with each passing decade regardless of their stage in the work-cycle. In contrast, the relationship between the employment share in highly routinary jobs and city population is positive. For example, in 2019 the share of high school workers in highly routinary jobs was 8 percentage points larger in the city with the largest population compared to the city with the smallest population. By contrast, the most populated city's employment share in low routine jobs is 4 percentage points smaller than the same share in the city with the smallest population. These patterns differ from the results of Autor (2019) who shows that the urban gradient in the employment share of highly skilled occupations is positive.[20] The difference arises because of the panel of occupations used. For example, Autor (2019) leaves out all non-farm occupations which have a very low RTI score and are usually performed in places with smaller population.

Second, workers in early stages of their career experienced different shifts in their occupational structure depending on their educational attainment. On the one hand, those with a college degree increased their employment share in jobs with a low RTI score over time. The increase fully absorbed the simultaneous decrease in highly routinary employment. In 2019 the average employment share in low routinary jobs is roughly ten percentage points larger than in 1980. During the same period, the employment in highly routinary jobs decreased 5 percentage points. On the other hand, the occupational structure for those with a high school diploma remained virtually unaltered.

Regardless of their educational attainment, workers in later stages of their work-cycle

[19] Autor and Dorn (2013) calculate the RTI score as $RTI = ln(R) - ln(M) - ln(A)$, where R, M, and A are, respectively, the routine, manual, and abstract task inputs in each occupation in 1980.

[20] Autor (2019) uses Autor and Dorn's (2013) definition of skill occupations. They catalogue a job as highly skilled if it was in the top tercile in the hourly wage distribution in 1980.

have shifted away from highly routinary jobs over time. The average employment share in jobs with a high RTI score was 23% in 1980, decreasing to 16% in 2019. However, the change in the occupational structure differs across places with different populations. In 1980 the fraction of high school workers performing highly routinary jobs was 13 percentage points higher in the most versus the least populated CBSA. In 2019 this difference was 8 percentage points. Conversely, workers with a college degree have historically had a lower fraction of employment in highly routinary jobs and a flattened urban gradient.

For old workers, the decrease in jobs with a high RTI score was fully absorbed by a simultaneous increase in low RTI employment. Both educational groups' increase fully absorbed the employment decrease in highly routinary jobs. Panels (e) and (f) of Figure 5 also show that the shift in the urban occupational structure differs between college and high school workers. In 1980 the fraction of college workers in a low RTI job was 5 percentage points lower in the most versus the least populated CBSA. In 2019 both fractions were roughly the same. For high school workers this gradient has changed little over time. From 1980 to 2019, the employment share in low RTI jobs was consistently 7 percentage points lower in the most versus the least populated city.

Overall, Figures 4 and 5 show that, through its effect on wages, the change in the occupational composition across age groups and cities explains the increase in both the level of the college wage gap and the gradient with respect to city size. These figures show that workers in early stages of their career saw college workers shift from low-paying-highly-routinary jobs to high-paying-low-routinary jobs, while high school graduates' occupational composition remained essentially unchanged over time. Also, workers in later stages of their career increased their share in high-paying-low-routinary jobs despite their educational attainment. However, the negative relationship between population size and the fraction of college workers performing low RTI jobs has eroded over time, causing an increase in the urban premia.

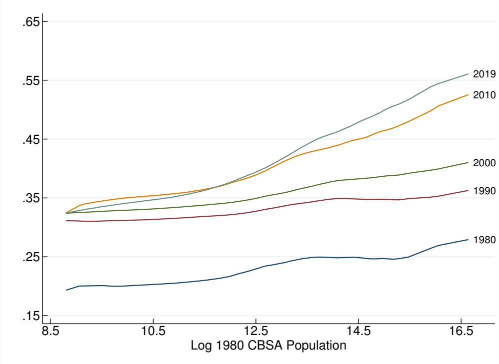
8 Conclusions

This paper argues that the rise in the U.S. spatial wage inequality over the last five decades has been driven by an increase in the college bias of agglomeration externalities for young workers. Using decennial population census and 5 years ACS, I estimate a flexible production function at the city level that allowed for imperfect substitution between workers of different ages and educational attainment. From the model I derive a simple equation linking changes in the college wage gap of an age group to the agglomeration externalities those workers experience. The parameters are then recovered using the exogenous variation in immigrants enclaves in 1970, a decade before the beginning of my database.

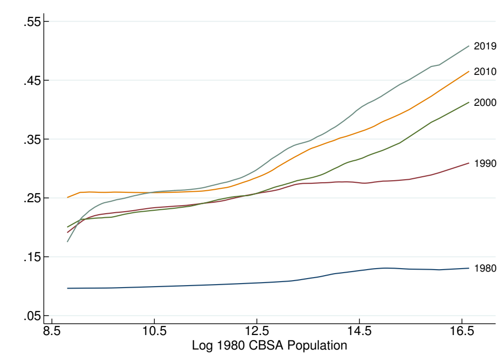
The results indicate that changes in the college bias of agglomeration externalities for young workers have been central to the increase in the wage inequality across local labours markets. The increase in the college wage gap for young workers located in the city with the largest population is 3.7% and 7.1% larger than the median city in 1980 and 2019, respectively. For old workers the same effect is 2.4% and 3.4% larger. City size accounts for about one third of the increase in wage inequality in the U.S. since 1980 (Baum-Snow and Pavan, 2013), so an important fraction of this effect can be traced back to increases in the college bias of agglomeration externalities through the work-cycle. My results also suggest that the college bias of agglomeration externalities through the work-cycle have increased because of shifts in the occupational structure. Younger workers with a high school diploma are trapped in low-paying highly-routinary jobs, whereas older workers with the same educational level have shifted away to less routinary jobs.

9 Figures

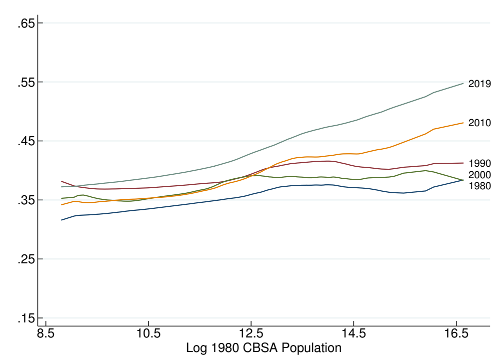
Figure 1: College to HS Graduates Wage Gap and City Size



(a) All ages



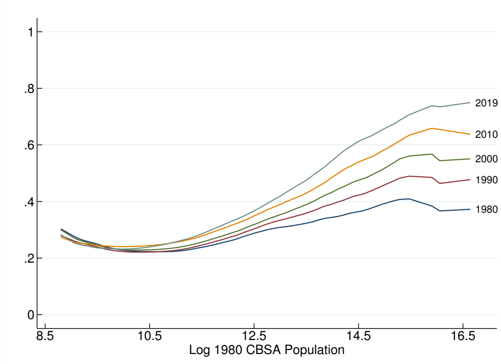
(b) 26-30 years old



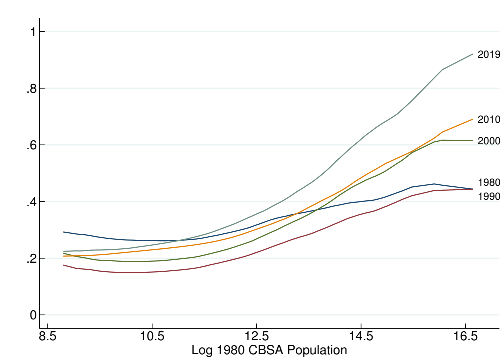
(c) 56-60 years old

Note: Full-time full-year men workers. Panel (a) presents the wage gap between college and high school graduates and its evolution with city size for the years 1980, 1990, 2000, 2010, and 2019. Panel (b) presents the same information but for the group of people that are between 26 and 30 years old at the moment of the census. Panel (c) does the same exercise for the group of people between 56 and 60 years old. The figures are in 2010 dollars. All values are smoothed using an Epanechnikov kernel.

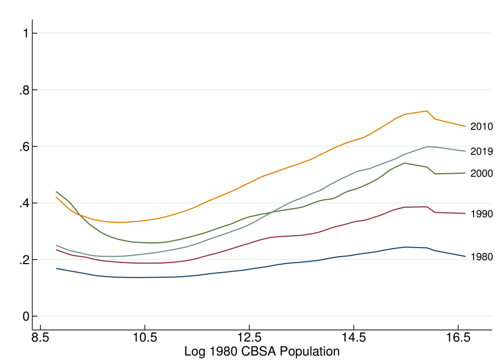
Figure 2: College to HS Relative Supply and City Size



(a) All ages



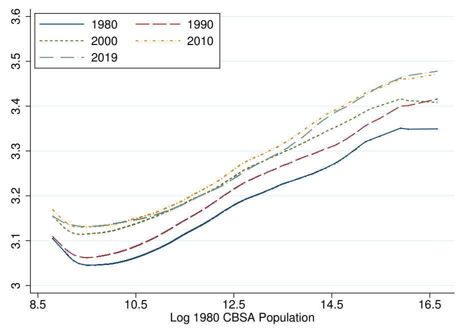
(b) 26-30 years old



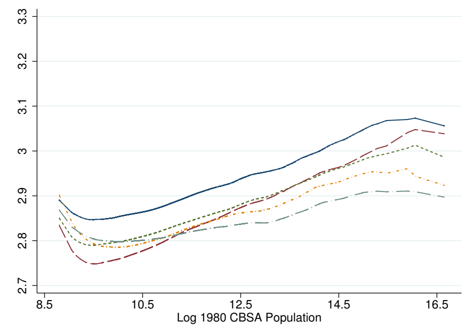
(c) 56-60 years old

Note: Full-time full-year men workers. Panel (a) presents the ratio of hours worked in a year between college and high school graduates and its evolution with city size for the years 1980, 1990, 2000, 2010, and 2019. Panel (b) presents the same information but for the group of people that are between 26 and 30 years old at the moment of the census. Panel (c) does the same exercise for the group of people between 56 and 60 years old. All values are smoothed using an Epanechnikov kernel.

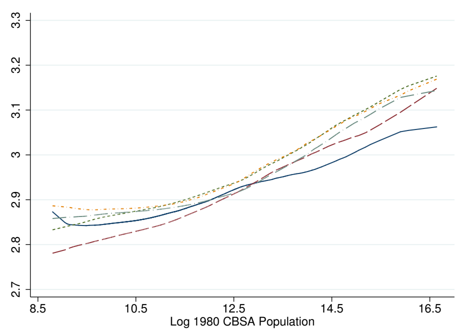
Figure 3: College and High School Graduates Wages and City Size



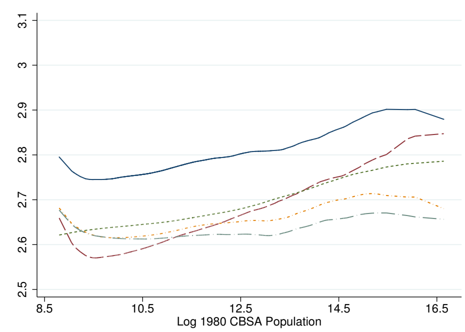
(a) College, All ages



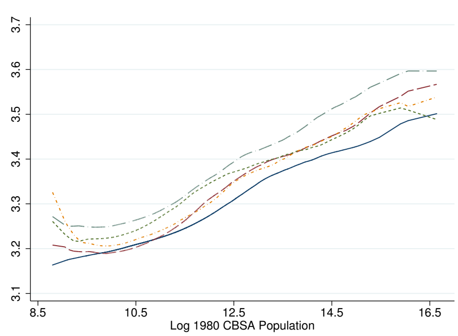
(b) High School, All ages



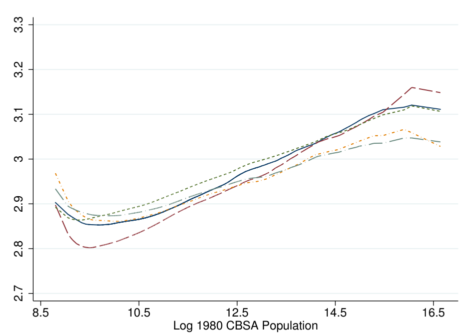
(c) College, 26-30 years old



(d) High School, 26-30 years old



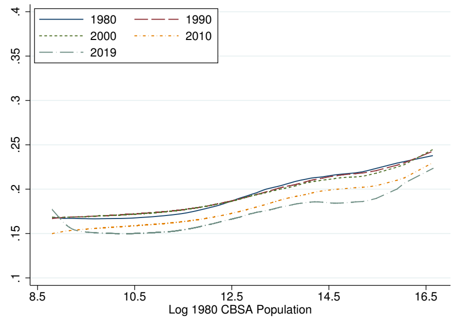
(e) College, 56-60 years old



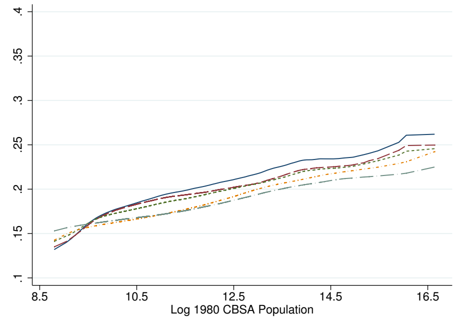
(f) High School, 56-60 years old

Note: Full-time full-year men workers. The left column presents the relationship between the average wage for college graduates and city size for all workers, those who were between 26-30, and those between 56-60. The right column presents the same information for high school graduates. The figures are in 2010 dollars. All values are smoothed using an Epanechnikov kernel.

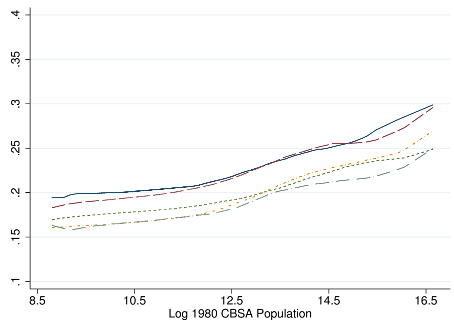
Figure 4: Share of Workers in High Routine Jobs and City Size



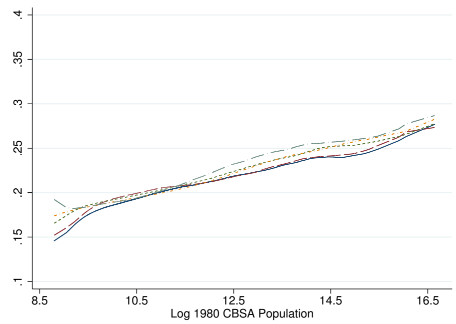
(a) College graduates, All ages



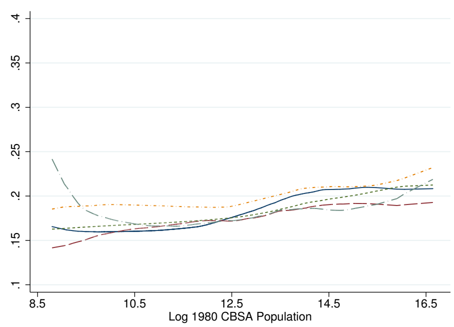
(b) High school graduates, All ages



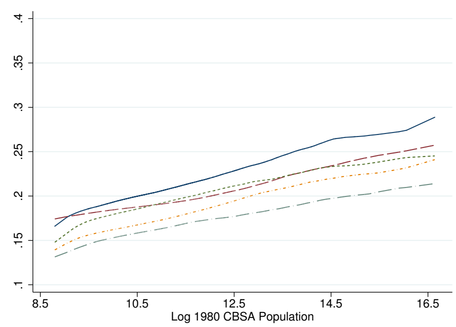
(c) College graduates, 26-30 years old



(d) High school graduates, 26-30 years old



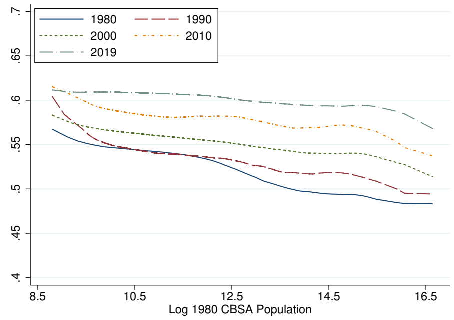
(e) College graduates, 56-60 years old



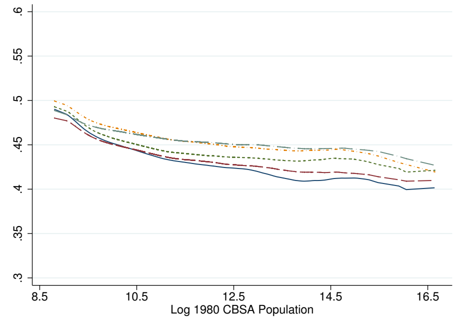
(f) High school graduates, 56-60 years old

Note: Full-time full-year men workers. The left column presents the relationship between the share of college workers performing highly routinely tasks and city size for all workers, those who were between 26-30, and those between 56-60. The right column presents the same information for high school graduates. For this figure I use the Routine Task Index (RTI) constructed by Autor and Dorn (2013). High routinary jobs are those who belong to the top tercile of the RTI distribution. All values are smoothed using an Epanechnikov kernel.

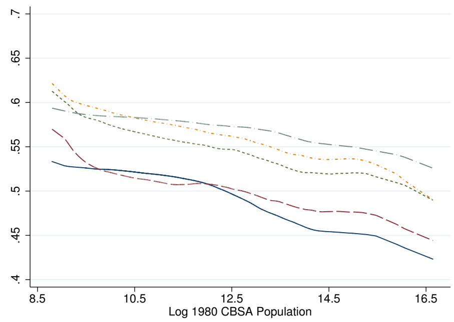
Figure 5: Share of Workers in Low Routine Jobs and City Size



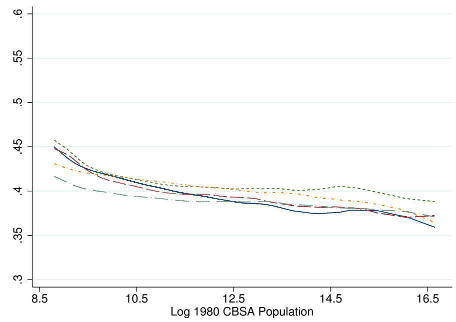
(a) College graduates, All ages



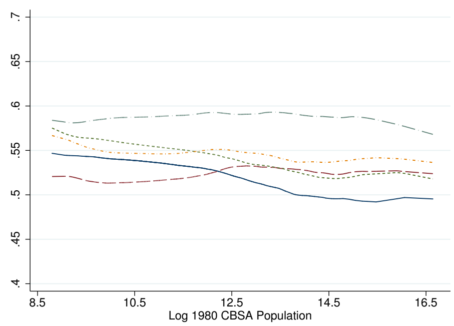
(b) High school graduates, All ages



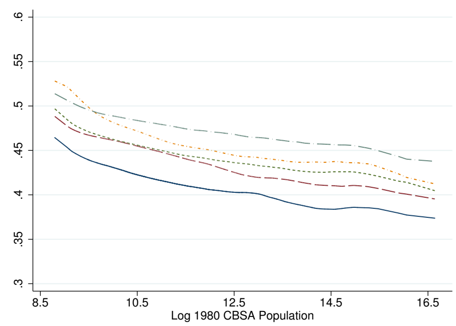
(c) College graduates, 26-30 years old



(d) High school graduates, 26-30 years old



(e) College graduates, 56-60 years old



(f) High school graduates, 56-60 years old

Note: Full-time full-year men workers. The left column presents the relationship between the share of college workers performing low routinary tasks and city size for all workers, those who were between 26-30, and those between 56-60. The right column presents the same information for high school graduates. For this figure I use the Routine Task Index (RTI) constructed by Autor and Dorn (2013). Low routinary jobs are those who belong to the bottom tercile of the RTI distribution. All values are smoothed using an Epanechnikov kernel.

10 Tables

Table 1: Patterns in log Wage Premia by Education, Age, and Location.

	College Wage Premium			Elasticity of Wages wrt 1980 City Population		
	All	Young	Old	All	Young	Old
1980	0.270[a]	0.139[a]	0.396[a]	0.044[a]	0.031[a]	0.048[a]
1990	0.377[a]	0.307[a]	0.442[a]	0.059[a]	0.058[a]	0.065[a]
2000	0.414[a]	0.356[a]	0.418[a]	0.051[a]	0.048[a]	0.051[a]
2010	0.493[a]	0.405[a]	0.465[a]	0.051[a]	0.043[a]	0.052[a]
2019	0.529[a]	0.435[a]	0.529[a]	0.052[a]	0.047[a]	0.053[a]

Note: All the coefficients belong to a different regression. [a] indicates the coef. is significant at the 1%, [b] at the 5%, and [c] at the 10%. Young refers to workers who are between 26-30 whereas old to those between 56-60. The first three columns correspond to a regression of (log) hourly wage and skill dummy. The definition of skill is college degree of more. The last three columns correspond to a regression of (log) hourly wage and 1980 city population. See Table A.1 for the table with standard errors. Calculations include census weights interacted with the labour supply for full-time full-year workers.

Table 2: Elasticities of College Wage Gap and Relative Factor Intensities wrt City Size.

	$\ln(w^s/w^u)$			ln(S/U)		
	All	Young	Old	All	Young	Old
1980	0.016[a]	0.014[a]	0.008[a]	0.109[a]	0.117[a]	0.118[a]
1990	0.009[a]	0.012[a]	0.006[a]	0.155[a]	0.225[a]	0.149[a]
2000	0.015[a]	0.033[a]	0.004[a]	0.179[a]	0.252[a]	0.148[a]
2010	0.033[a]	0.044[a]	0.025[a]	0.195[a]	0.254[a]	0.144[a]
2019	0.043[a]	0.048[a]	0.030[a]	0.224[a]	0.292[a]	0.209[a]

Note: All the coefficients belong to a different regression. ^{a} indicates the coef. is significant at the 1%, ^{b} at the 5%, and ^{c} at the 10%. The first three columns show the effect city size has on the (log) ratio between average wage for college and non-college workers. The last three columns show the effect city size has on the (log) ratio between total hours worked by college and non-college workers. The definition of skill is college degree of more. See Table A.2 for the table with standard errors. Regressions are weighted by 1980 city population.

Table 3: Change in the Wage Gap and Quantities with Respect to City Size.

	All		Young		Old	
	Δ ln(w^s/w^u)	Δ ln(S/U)	Δ ln(w^s/w^u)	Δ ln(S/U)	Δ ln(w^s/w^u)	Δ ln(S/U)
ln(1980 City Population)	0.007[a]	0.027[a]	0.008[a]	0.042[a]	0.006[a]	0.020[a]
	(0.000)	(0.001)	(0.001)	(0.003)	(0.001)	(0.003)
1990 - 2000 Indicator	-0.062[a]	-0.039[a]	-0.105[a]	0.473[a]	-0.056[a]	-0.122[a]
	(0.002)	(0.006)	(0.005)	(0.015)	(0.007)	(0.016)
2000 - 2010 Indicator	-0.032[a]	-0.017[a]	-0.094[a]	0.344[a]	0.007[a]	-0.102[a]
	(0.002)	(0.006)	(0.005)	(0.015)	(0.007)	(0.016)
2010 - 2019 Indicator	-0.072[a]	-0.048[a]	-0.128[a]	0.399[a]	0.019[a]	-0.715[a]
	(0.002)	(0.006)	(0.005)	(0.015)	(0.007)	(0.016)
Constant	0.001	-0.249[a]	0.028[a]	-0.811[a]	-0.042[a]	0.154[a]
	(0.007)	(0.018)	(0.014)	(0.041)	(0.019)	(0.046)
Observations	3,668	3,668	3,668	3,668	3,666	3,666
R^2	0.246	0.125	0.192	0.286	0.042	0.393

Note: [a] indicates the coef. is significant at the 1%, [b] at the 5%, and [c] at the 10%. Each column reports the coefficients and standard errors of a different regression. The definition of skill is college degree of more. Regressions are weighted by 1980 city population.

Table 4: Supply Shock Regression by Education.

	Δ ln(S/U)	
	Age	All
Δ Predicted ln(S/U)	0.699[a]	0.097[a]
	(0.036)	(0.032)
ln(1980 City Population)	0.089[a]	0.659[a]
	(0.008)	(0.039)
$\ln(S/U)_{t-1}$	-0.372[a]	-0.175[a]
	(0.030)	(0.013)
Observations	25,672	3,668
R^2	0.364	0.111

Note: [a] indicates the coef. is significant at the 1%, [b] at the 5%, and [c] at the 10%. The definition of skill is college degree of more. All regression controlled for year fixed effects. First column also controls for age group fixed effect. Regressions are weighted by 1980 city population.

Table 5: IV Regression Results.

	Dep. var.: City-age Wage Gap			
	(1)	(2)	(3)	(4)
$\Delta\ \log\left(\frac{S_{ct}}{U_{ct}}\right)$	-0.030	-0.028	-0.057	-0.045
	(0.093)	(0.113)	(0.097)	(0.098)
$\Delta\ \log\left(\frac{S_{ctj}}{U_{ctj}}\right)$	-0.230$^{\mathrm{a}}$	-0.234$^{\mathrm{a}}$	-0.218$^{\mathrm{a}}$	-0.229$^{\mathrm{a}}$
	(0.042)	(0.044)	(0.041)	(0.043)
$\log\left(\frac{S_{ct}^{imm}}{U_{ct}^{imm}}\right)$		0.003	0.003	0.003
		(0.003)	(0.003)	(0.003)
Implied σ_A	4.348$^{\mathrm{a}}$	4.282$^{\mathrm{a}}$	4.597$^{\mathrm{a}}$	4.363$^{\mathrm{a}}$
	(0.803)	(0.815)	(0.862)	(0.810)
Implied σ_E	3.844$^{\mathrm{a}}$	3.821$^{\mathrm{a}}$	3.648$^{\mathrm{a}}$	3.648$^{\mathrm{a}}$
	(1.007)	(1.243)	(0.998)	(0.998)
Time FE	Yes	No	Yes	Yes
Age FE	Yes	Yes	No	Yes
Time Trend	No	Yes	No	No
Observations	25,672	25,560	25,560	25,560
First-Stage F-statistic	27.438	33.671	27.937	27.220

Note: ^{a} indicates the coef. is significant at the 1%, ^{b} at the 5%, and ^{c} at the 10%. Each column presents the IV estimates of Equation (8). Shift-share instruments are used for the change in the log supply of skilled to unskilled workers. More information can be found in Section 4.3. Regressions are weighted by 1980 city population and standard errors are clustered at the CBSA level.

Table 6: Agglomeration Externalities by Age Group and Year.

	Age Range							Age
	26-30	31-35	36-40	41-45	46-50	51-55	56-60	p-val
1990	0.012[b]	0.010[c]	0.010[c]	0.010	0.005	0.004	0.008	0.000
	(0.006)	(0.006)	(0.006)	(0.006)	(0.006)	(0.007)	(0.006)	
2000	0.019[a]	0.018[a]	0.011[a]	0.007[b]	0.008[a]	0.011[a]	0.008[a]	0.000
	(0.002)	(0.002)	(0.003)	(0.004)	(0.003)	(0.003)	(0.003)	
2010	0.025[a]	0.028[a]	0.027[a]	0.023[a]	0.017[a]	0.016[a]	0.021[a]	0.000
	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	
2019	0.023[a]	0.023[a]	0.020[a]	0.020[a]	0.019[a]	0.016[a]	0.011[a]	0.000
	(0.003)	(0.002)	(0.003)	(0.003)	(0.003)	(0.004)	(0.003)	
Year p-val	0.017	0.006	0.000	0.000	0.000	0.005	0.013	

Note: ^{a} indicates the coefficient is significant at the 1%, ^{b} at the 5%, and ^{c} at the 10%. All the coefficients come from Table 5, Column 4. The last column presents the p-value that comes from testing for each decade if the effect between age groups is constant. The last row presents the p-values that comes from testing if for every age group the effect is constant across decades.

References

Acemoglu, Daron and Pascual Restrepo, "Tasks, Automation, and the Rise in US Wage Inequality," Working Paper 28920, National Bureau of Economic Research June 2021.

Autor, David, *Work of the Past, Work of the Future*, Vol. 109, National Bureau of Economic Research Cambridge, MA, 2019.

_ **and David Dorn**, "The growth of low-skill service jobs and the polarization of the US labor market," *American Economic Review*, 2013, *103* (5), 1553–97.

Autor, David H., Frank Levy, and Richard J. Murnane, "The Skill Content of Recent Technological Change: An Empirical Exploration," *The Quarterly Journal of Economics*, 11 2003, *118* (4), 1279–1333.

Autor, David H, Lawrence F Katz, and Melissa S Kearney, "Trends in US wage inequality: Revising the revisionists," *The Review of economics and statistics*, 2008, *90* (2), 300–323.

Azar, José, Ioana Marinescu, Marshall Steinbaum, and Bledi Taska, "Concentration in US labor markets: Evidence from online vacancy data," *Labour Economics*, 2020, *66*, 101886.

Baum-Snow, Nathaniel and Ronni Pavan, "Inequality and city size," *Review of Economics and Statistics*, 2013, *95* (5), 1535–1548.

_ **, Matthew Freedman, and Ronni Pavan**, "Why has urban inequality increased?," *American Economic Journal: Applied Economics*, 2018, *10* (4), 1–42.

Borjas, George J., "The Labor Demand Curve is Downward Sloping: Reexamining the Impact of Immigration on the Labor Market*," *The Quarterly Journal of Economics*, 11 2003, *118* (4), 1335–1374.

Borjas, George J, Jeffrey Grogger, and Gordon H Hanson, "Imperfect Substitution between Immigrants and Natives: A Reappraisal," Working Paper 13887, National Bureau of Economic Research March 2008.

Bound, John and George Johnson, "Changes in the Structure of Wages in the 1980's: An Evaluation of Alternative Explanations," *The American Economic Review*, 1992, *82* (3), 371–392.

Card, David, "Immigrant inflows, native outflows, and the local labor market impacts of higher immigration," *Journal of Labor Economics*, 2001, *19* (1), 22–64.

_ , "Immigration and Inequality," *American Economic Review*, May 2009, *99* (2), 1–21.

_ **and John DiNardo**, "Do Immigrant Inflows Lead to Native Outflows?," *American Economic Review*, May 2000, *90* (2), 360–367.

_ **and Thomas Lemieux**, "Can falling supply explain the rising return to college for younger men? A cohort-based analysis," *The Quarterly Journal of Economics*, 2001, *116* (2), 705–746.

Ciccone, Antonio and Giovanni Peri, "Long-Run Substitutability Between More and Less Educated Workers: Evidence from U.S. States, 1950–1990," *The Review of Economics and Statistics*, 11 2005, *87* (4), 652–663.

Combes, Pierre-Philippe and Laurent Gobillon, "The empirics of agglomeration economies," in "Handbook of regional and urban economics," Vol. 5, Elsevier, 2015, pp. 247–348.

_ , Gilles Duranton, and Laurent Gobillon, "Spatial wage disparities: Sorting matters!," *Journal of urban economics*, 2008, *63* (2), 723–742.

Costa, Dora L. and Matthew E. Kahn, "Power Couples: Changes in the Locational Choice of the College Educated, 1940–1990*," *The Quarterly Journal of Economics*, 11 2000, *115* (4), 1287–1315.

Davis, Donald R, Eric Mengus, and Tomasz K Michalski, "Labor Market Polarization and The Great Divergence: Theory and Evidence," Technical Report, National Bureau of Economic Research 2020.

Diamond, Rebecca, "The determinants and welfare implications of US workers' diverging location choices by skill: 1980-2000," *American Economic Review*, 2016, *106* (3), 479–524.

Duranton, Gilles and Diego Puga, "Urban Growth and its Aggregate Implications," Working Paper 26591, National Bureau of Economic Research December 2019.

Dustmann, Christian, Tommaso Frattini, and Ian P. Preston, "The Effect of Immigration along the Distribution of Wages," *The Review of Economic Studies*, 04 2012, *80* (1), 145–173.

Eckert, Fabian, Sharat Ganapati, and Conor Walsh, "Skilled tradable services: The transformation of US high-skill labor markets," *Available at SSRN 3439118*, 2019.

Florida, Richard, *The rise of the creative class*, Vol. 9, Basic books New York, 2002.

Giannone, Elisa, "Skilled-biased technical change and regional convergence," *University of Chicago. Unpublished manuscript*, 2017.

Glaeser, Edward L and David C Mare, "Cities and skills," *Journal of labor economics*, 2001, *19* (2), 316–342.

Goldsmith-Pinkham, Paul, Isaac Sorkin, and Henry Swift, "Bartik instruments: What, when, why, and how," *American Economic Review*, 2020, *110* (8), 2586–2624.

Gould, E. D., "Cities, Workers, and Wages: A Structural Analysis of the Urban Wage Premium," *The Review of Economic Studies*, 04 2007, *74* (2), 477–506.

Jaeger, David A, Joakim Ruist, and Jan Stuhler, "Shift-Share Instruments and the Impact of Immigration," Working Paper 24285, National Bureau of Economic Research February 2018.

Katz, Lawrence F and Kevin M Murphy, "Changes in relative wages, 1963–1987: supply and demand factors," *The quarterly journal of economics*, 1992, *107* (1), 35–78.

Lee, Yongsung, Bumsoo Lee, and Md Tanvir Hossain Shubho, "Urban revival by Millennials? Intraurban net migration patterns of young adults, 1980–2010," *Journal of Regional Science*, 2019, *59* (3), 538–566.

Lewis, Ethan, "Immigration, Skill Mix, and Capital Skill Complementarity*," *The Quarterly Journal of Economics*, 05 2011, *126* (2), 1029–1069.

Lindley, Joanne and Stephen Machin, "Spatial changes in labour market inequality," *Journal of Urban Economics*, 2014, *79*, 121–138. Spatial Dimensions of Labor Markets.

Manacorda, Marco, Alan Manning, and Jonathan Wadsworth, "The Impact of Immigration on the Structure of Male Wages: Theory and Evidence from Britain," CReAM Discussion Paper Series 0608, Centre for Research and Analysis of Migration (CReAM), Department of Economics, University College London 2006.

Moretti, Enrico, *The new geography of jobs*, Houghton Mifflin Harcourt, 2012.

_ , "Real wage inequality," *American Economic Journal: Applied Economics*, 2013, *5* (1), 65–103.

Ottaviano, Gianmarco I. P. and Giovanni Peri, "Rethinking the Effect of Immigration on Wages," *Journal of the European Economic Association*, 02 2012, *10* (1), 152–197.

Roca, Jorge De La and Diego Puga, "Learning by working in big cities," *The Review of Economic Studies*, 2017, *84* (1), 106–142.

Ruggles, Steven, Sarah Flood, Ronald Goeken, Josiah Grover, Erin Meyer, Jose Pacas, and Matthew Sobek, "IPUMS USA: Version 10.0 [dataset].," Minneapolis, MN: IPUMS. https://doi.org/10.18128/D010.V10.0 2020.

Wheeler, Christopher H, "Cities and the growth of wages among young workers: Evidence from the NLSY," *Journal of Urban Economics*, 2006, *60* (2), 162–184.

A Tables Appendix

A.1 Motivational Patterns

Table A.1: Patterns in log Wage Premia by Education, Age, and Location.

	College Wage Premium			Elasticity of Wages wrt 1980 City Population		
	All	Young	Old	All	Young	Old
1980	0.270^a	0.139^a	0.396^a	0.044^a	0.031^a	0.048^a
	(0.001)	(0.001)	(0.002)	(0.000)	(0.000)	(0.000)
1990	0.377^a	0.307^a	0.442^a	0.059^a	0.058^a	0.065^a
	(0.001)	(0.001)	(0.002)	(0.000)	(0.000)	(0.001)
2000	0.414^a	0.356^a	0.418^a	0.051^a	0.048^a	0.051^a
	(0.000)	(0.001)	(0.002)	(0.000)	(0.000)	(0.000)
2010	0.493^a	0.405^a	0.465^a	0.051^a	0.043^a	0.052^a
	(0.000)	(0.001)	(0.001)	(0.000)	(0.000)	(0.000)
2019	0.529^a	0.435^a	0.529^a	0.052^a	0.047^a	0.053^a
	(0.000)	(0.001)	(0.001)	(0.000)	(0.000)	(0.000)

Note: All the coefficients belong to a different regression. ^{a} indicates the coef. is significant at the 1%, ^{b} at the 5%, and ^{c} at the 10%. Young refers to workers who are between 26-30 whereas old to those between 56-60. The first three columns correspond to a regression of (log) hourly wage and skill dummy. The definition of skill is college degree of more. The last three columns correspond to a regression of (log) hourly wage and 1980 city population. Calculations include census weights interacted with the labour supply for full-time full-year workers.

Table A.2: Elasticities of College Wage Gap and Relative Factor Intensities wrt City Size.

	w^s/w^u			S/U		
	All	Young	Old	All	Young	Old
1980	0.016[a]	0.014[a]	0.008[a]	0.109[a]	0.117[a]	0.118[a]
	(0.001)	(0.002)	(0.003)	(0.005)	(0.005)	(0.008)
1990	0.009[a]	0.012[a]	0.006[a]	0.155[a]	0.225[a]	0.149[a]
	(0.001)	(0.002)	(0.003)	(0.005)	(0.006)	(0.008)
2000	0.015[a]	0.033[a]	0.004[a]	0.179[a]	0.252[a]	0.148[a]
	(0.001)	(0.002)	(0.002)	(0.005)	(0.007)	(0.007)
2010	0.033[a]	0.044[a]	0.025[a]	0.195[a]	0.254[a]	0.144[a]
	(0.001)	(0.002)	(0.002)	(0.005)	(0.008)	(0.006)
2019	0.043[a]	0.048[a]	0.030[a]	0.224[a]	0.292[a]	0.209[a]
	(0.001)	(0.002)	(0.002)	(0.006)	(0.008)	(0.007)

Note: All the coefficients belong to a different regression. [a] indicates the coef. is significant at the 1%, [b] at the 5%, and [c] at the 10%. The first three columns show the effect city size has on the (log) ratio between average wage for college and non-college workers. The last three columns show the effect city size has on the (log) ratio between total hours worked by college and non-college workers. The definition of skill is college degree of more.Regressions are weighted by 1980 city population.

A.2 Main Results

In this section I present the fist stage estimates for the preferred specification and the estimates for the college bias agglomeration externalities for the specification 1-3 in Table 5.

Table A.3: First Stage

	Δ ln(S/U)	Δ ln(S_j/U_j)
$\Delta\ ln(\hat{S}/U)$	0.191[a]	-0.155[a]
	(0.047)	(0.069)
$\Delta\ ln(\hat{S}/U)_j$	0.056[a]	0.431[a]
	(0.016)	(0.065)
$ln(S/U)_{t-1}$	0.000	0.002
	(0.005)	(0.005)
Observations	25,560	25,560
R^2	0.181	0.367
F test excl. instr.	19.312	32.014

Note: ^{a} indicates the coef. is significant at the 1%, ^{b} at the 5%, and ^{c} at the 10%. All regression control for year and age fixed effects. Also, they control for the effect of agglomeration externalities across age groups and years. Regressions are weighted by 1980 city population.

Table A.4: Agglomeration Externalities by Age Group and Year.
No previous immigration.

	Age Range							Age
	26-30	31-35	36-40	41-45	46-50	51-55	56-60	p-val
1990	0.011^{b}	0.009^{c}	0.009	0.009	0.004	0.003	0.007	0.000
	(0.006)	(0.006)	(0.006)	(0.006)	(0.006)	(0.007)	(0.006)	
2000	0.019^{a}	0.018^{a}	0.011^{a}	0.007^{c}	0.008^{a}	0.011^{a}	0.008^{a}	0.000
	(0.002)	(0.002)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	
2010	0.025^{a}	0.028^{a}	0.027^{a}	0.023^{a}	0.017^{a}	0.016^{a}	0.021^{a}	0.000
	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	
2019	0.023^{a}	0.023^{a}	0.019^{a}	0.020^{a}	0.019^{a}	0.016^{a}	0.011^{a}	0.000
	(0.003)	(0.002)	(0.003)	(0.003)	(0.003)	(0.004)	(0.003)	
Year p-val	0.008	0.003	0.000	0.000	0.000	0.002	0.010	

Note: ^{a} indicates the coefficient is significant at the 1%, ^{b} at the 5%, and ^{c} at the 10%. All the coefficients come from Table 5, Column 1. The last column presents the p-value that comes from testing for each decade if the effect between age groups is constant. The last row presents the p-values that comes from testing if for every age group the effect is constant across decades.

Table A.5: Agglomeration Externalities by Age Group and Year.
Linear Time Trend.

	Age Range							Age
	26-30	31-35	36-40	41-45	46-50	51-55	56-60	p-val
1990	0.013[b]	0.011[b]	0.011[b]	0.010[c]	0.006	0.005	0.009	0.000
	(0.005)	(0.005)	(0.005)	(0.006)	(0.006)	(0.006)	(0.005)	
2000	0.018[a]	0.017[a]	0.010[b]	0.006	0.007[c]	0.010[a]	0.007[c]	0.000
	(0.003)	(0.003)	(0.004)	(0.005)	(0.004)	(0.004)	(0.004)	
2010	0.022[a]	0.025[a]	0.024[a]	0.020[a]	0.014[a]	0.013[a]	0.018[a]	0.000
	(0.002)	(0.002)	(0.002)	(0.003)	(0.003)	(0.003)	(0.002)	
2019	0.025[a]	0.025[a]	0.021[a]	0.022[a]	0.021[a]	0.018[a]	0.013[a]	0.000
	(0.003)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.003)	
Year p-val	0.133	0.000	0.000	0.000	0.000	0.000	0.000	

Note: [a] indicates the coefficient is significant at the 1%, [b] at the 5%, and [c] at the 10%. All the coefficients come from Table 5, Column 2. The last column presents the p-value that comes from testing for each decade if the effect between age groups is constant. The last row presents the p-values that comes from testing if for every age group the effect is constant across decades.

Table A.6: Agglomeration Externalities by Age Group and Year.
No Age FE.

	Age Range							Age
	26-30	31-35	36-40	41-45	46-50	51-55	56-60	p-val
1990	0.009	0.006	0.008	0.010^{c}	0.008	0.007	0.011^{c}	0.000
	(0.006)	(0.006)	(0.006)	(0.006)	(0.006)	(0.006)	(0.006)	
2000	0.015^{a}	0.014^{a}	0.010^{a}	0.008^{b}	0.011^{a}	0.014^{a}	0.012^{a}	0.000
	(0.002)	(0.003)	(0.003)	(0.003)	(0.003)	(0.002)	(0.003)	
2010	0.021^{a}	0.024^{a}	0.025^{a}	0.024^{a}	0.020^{a}	0.019^{a}	0.024^{a}	0.000
	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	
2019	0.019^{a}	0.018^{a}	0.018^{a}	0.021^{a}	0.022^{a}	0.019^{a}	0.015^{a}	0.000
	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	
Year p-val	0.019	0.007	0.000	0.000	0.000	0.003	0.012	

Note: ^{a} indicates the coefficient is significant at the 1%, ^{b} at the 5%, and ^{c} at the 10%. All the coefficients come from Table 5, Column 3. The last column presents the p-value that comes from testing for each decade if the effect between age groups is constant. The last row presents the p-values that comes from testing if for every age group the effect is constant across decades.

A.2.1 Main Results: Sorting.

Table A.7: Agglomeration Externalities by Age Group and Year.
Testing Sorting.

	Age Range							Age
	26-30	31-35	36-40	41-45	46-50	51-55	56-60	p-val
1990	0.004	0.003	-0.001	-0.004	-0.005	-0.004	-0.001	0.000
	(0.006)	(0.006)	(0.005)	(0.005)	(0.006)	(0.005)	(0.005)	
2000	0.005[a]	0.006[a]	0.004	0.001	0.001	0.001	-0.001	0.000
	(0.002)	(0.002)	(0.003)	(0.003)	(0.002)	(0.002)	(0.002)	
2010	0.014[a]	0.014[a]	0.014[a]	0.012[a]	0.011[a]	0.013[a]	0.013[a]	0.848
	(0.002)	(0.002)	(0.003)	(0.002)	(0.003)	(0.003)	(0.003)	
2019	0.014[a]	0.013[a]	0.010[a]	0.010[a]	0.010[a]	0.012[a]	0.010[a]	0.001
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.003)	(0.003)	
Year p-val	0.000	0.000	0.008	0.000	0.000	0.000	0.000	

Note: [a] indicates the coefficient is significant at the 1%, [b] at the 5%, and [c] at the 10%. All the coefficients come from the same regression. The last column presents the p-value of testing for each decade if the effect between age groups is constant. The last row presents the p-values of testing if for every age group the effect is constant across decades.The F-Statistic from the first stage is 17.417.

A.2.2 Main Results: Calibrated Parameters.

In this section I use estimates that come from the labour literature for the elasticity of substitution between workers with and without college degree and between workers of different ages.

Table A.8: Agglomeration Externalities by Age Group and Year.
$\sigma_A = 4.5$ and $\sigma_E = 1.4$.

	Age Range							Age
	26-30	31-35	36-40	41-45	46-50	51-55	56-60	p-val
1990	0.033^{a}	0.031^{a}	0.031^{a}	0.030^{a}	0.026^{a}	0.025^{a}	0.029^{a}	0.000
	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.004)	(0.003)	
2000	0.029^{a}	0.029^{a}	0.022^{a}	0.018^{a}	0.019^{a}	0.022^{a}	0.019^{a}	0.000
	(0.003)	(0.003)	(0.003)	(0.002)	(0.002)	(0.003)	(0.002)	
2010	0.032^{a}	0.036^{a}	0.035^{a}	0.031^{a}	0.025^{a}	0.024^{a}	0.028^{a}	0.000
	(0.005)	(0.005)	(0.005)	(0.004)	(0.004)	(0.004)	(0.004)	
2019	0.035^{a}	0.035^{a}	0.032^{a}	0.033^{a}	0.031^{a}	0.029^{a}	0.024^{a}	0.000
	(0.003)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	
Year p-val	0.282	0.191	0.003	0.000	0.000	0.077	0.028	

Note: ^{a} indicates the coefficient is significant at the 1%, ^{b} at the 5%, and ^{c} at the 10%. All the coefficients come from regression that calibrates $\sigma_A = 4.5$ and $\sigma_E = 1.4$. The last column presents the p-value that comes from testing for each decade if the effect between age groups is constant. The last row presents the p-values that comes from testing if for every age group the effect is constant across decades.

Table A.9: Agglomeration Externalities by Age Group and Year.
$\sigma_A = 4.5$ and $\sigma_E = 1.7$.

	Age Range							Age
	26-30	31-35	36-40	41-45	46-50	51-55	56-60	p-val
1990	0.027[a]	0.025[a]	0.025[a]	0.024[a]	0.020[a]	0.019[a]	0.023[a]	0.000
	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.004)	(0.003)	
2000	0.026[a]	0.026[a]	0.019[a]	0.015[a]	0.016[a]	0.019[a]	0.016[a]	0.000
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	
2010	0.030[a]	0.033[a]	0.033[a]	0.028[a]	0.023[a]	0.022[a]	0.026[a]	0.000
	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.003)	(0.004)	
2019	0.031[a]	0.031[a]	0.028[a]	0.029[a]	0.027[a]	0.025[a]	0.020[a]	0.000
	(0.003)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	
Year p-val	0.201	0.102	0.001	0.000	0.000	0.032	0.042	

Note: ^{a} indicates the coefficient is significant at the 1%, ^{b} at the 5%, and ^{c} at the 10%. All the coefficients come from regression that calibrates $\sigma_A = 4.5$ and $\sigma_E = 1.7$. The last column presents the p-value that comes from testing for each decade if the effect between age groups is constant. The last row presents the p-values that comes from testing if for every age group the effect is constant across decades.

Table A.10: Agglomeration Externalities by Age Group and Year.
$\sigma_A = 5.5$ and $\sigma_E = 1.4$.

	Age Range							Age
	26-30	31-35	36-40	41-45	46-50	51-55	56-60	p-val
1990	0.033^{a}	0.032^{a}	0.031^{a}	0.030^{a}	0.026^{a}	0.025^{a}	0.029^{a}	0.000
	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	
2000	0.028^{a}	0.028^{a}	0.023^{a}	0.019^{a}	0.019^{a}	0.021^{a}	0.019^{a}	0.000
	(0.003)	(0.003)	(0.003)	(0.002)	(0.002)	(0.003)	(0.002)	
2010	0.031^{a}	0.035^{a}	0.034^{a}	0.031^{a}	0.026^{a}	0.025^{a}	0.028^{a}	0.000
	(0.005)	(0.005)	(0.005)	(0.004)	(0.004)	(0.004)	(0.004)	
2019	0.034^{a}	0.034^{a}	0.032^{a}	0.033^{a}	0.031^{a}	0.029^{a}	0.025^{a}	0.000
	(0.003)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	
Year p-val	0.213	0.187	0.012	0.000	0.000	0.024	0.022	

Note: ^{a} indicates the coefficient is significant at the 1%, ^{b} at the 5%, and ^{c} at the 10%. All the coefficients come from regression that calibrates $\sigma_A = 5.5$ and $\sigma_E = 1.4$. The last column presents the p-value that comes from testing for each decade if the effect between age groups is constant. The last row presents the p-values that comes from testing if for every age group the effect is constant across decades.

Table A.11: Agglomeration Externalities by Age Group and Year.
$\sigma_A = 5.5$ and $\sigma_E = 1.7$.

	Age Range							Age
	26-30	31-35	36-40	41-45	46-50	51-55	56-60	p-val
1990	0.028[a]	0.026[a]	0.025[a]	0.024[a]	0.020[a]	0.019[a]	0.023[a]	0.000
	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.004)	(0.003)	
2000	0.025[a]	0.025[a]	0.020[a]	0.016[a]	0.016[a]	0.018[a]	0.016[a]	0.000
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	
2010	0.029[a]	0.032[a]	0.032[a]	0.029[a]	0.024[a]	0.023[a]	0.026[a]	0.000
	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.003)	(0.004)	
2019	0.030[a]	0.031[a]	0.028[a]	0.029[a]	0.028[a]	0.026[a]	0.022[a]	0.000
	(0.003)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	
Year p-val	0.200	0.132	0.006	0.000	0.000	0.007	0.028	

Note: ^{a} indicates the coefficient is significant at the 1%, ^{b} at the 5%, and ^{c} at the 10%. All the coefficients come from regression that calibrates $\sigma_A = 5.5$ and $\sigma_E = 1.7$. The last column presents the p-value that comes from testing for each decade if the effect between age groups is constant. The last row presents the p-values that comes from testing if for every age group the effect is constant across decades.

A.2.3 Main Results: Efficiency Units.

Table A.12: Agglomeration Externalities by Age Group and Year.
Efficiency Units.

	Age Range							Age
	26-30	31-35	36-40	41-45	46-50	51-55	56-60	p-val
1990	0.003	0.002	0.003	0.004	-0.000	-0.001	0.003	0.000
	(0.006)	(0.006)	(0.005)	(0.005)	(0.006)	(0.006)	(0.005)	
2000	0.010^{a}	0.011^{a}	0.005^{c}	0.002	0.003	0.007^{a}	0.004^{b}	0.000
	(0.002)	(0.002)	(0.003)	(0.003)	(0.002)	(0.002)	(0.002)	
2010	0.014^{a}	0.018^{a}	0.017^{a}	0.014^{a}	0.008^{a}	0.008^{a}	0.012^{a}	0.000
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	
2019	0.017^{a}	0.018^{a}	0.015^{a}	0.016^{a}	0.014^{a}	0.011^{a}	0.006^{b}	0.000
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.003)	(0.003)	
Year p-val	0.000	0.000	0.000	0.000	0.000	0.000	0.053	

Note: ^{a} indicates the coefficient is significant at the 1%, ^{b} at the 5%, and ^{c} at the 10%. All the coefficients come from the same regression. Wages and quantities are in Efficiency Units. The F-statistic of the first stage is 25.677. The last column presents the p-value that comes from testing for each decade if the effect between age groups is constant. The last row presents the p-values that comes from testing if for every age group the effect is constant across decades.

A.2.4 Main Results: Only Big Cities.

In this section I use the subsample of cities that are 'big'. I use four different thresholds to classify a city as big. I use those cities that had 25,000, 50,000, 75,000 and 100,000 or more inhabitants in 1980.

Table A.13: IV Regression Results.
Big Cities.

	Dep. var.: City-age Wage Gap			
	(1)	(2)	(3)	(4)
$\Delta\ \log\left(\frac{S_{ct}}{U_{ct}}\right)$	-0.041	-0.030	-0.046	-0.050
	(0.100)	(0.102)	(0.109)	(0.111)
$\Delta\ \log\left(\frac{S_{ctj}}{U_{ctj}}\right)$	-0.234^{a}	-0.256^{a}	-0.259^{a}	-0.264^{a}
	(0.044)	(0.047)	(0.051)	(0.052)
$\log\left(\frac{S^{imm}_{ct}}{U^{imm}_{ct}}\right)$	0.003	0.005	0.007^{c}	0.008^{b}
	(0.003)	(0.003)	(0.004)	(0.004)
Implied σ_A	4.273^{a}	3.904^{a}	3.864^{a}	3.795^{a}
	(0.799)	(0.717)	(0.758)	(0.748)
Implied σ_E	3.632^{a}	3.496^{a}	3.285^{a}	3.191^{a}
	(0.994)	(0.928)	(0.876)	(0.835)
Observations	22,410	14,810	11,067	8,554
First-Stage F-statistic	26.026	23.819	22.800	23.886

Note: ^{a} indicates the coef. is significant at the 1%, ^{b} at the 5%, and ^{c} at the 10%. Each column presents the IV estimates of Equation (8). Shift-share instruments are used for the change in the log supply of skilled to unskilled workers. More information can be found in Section 4.3. Column 1 shows the results using the subsample of cities with a population higher than 25,000 inhabitants in 1980 (804 cities). Column 2 those with more than 50,000 (530). Column 3 those with more than 75,000 (396). Column 4 those with more than 100,000 (306). All regressions includes time and age group fixed effects. Regressions are weighted by 1980 city population and standard errors are clustered at the CBSA level.

Table A.14: Agglomeration Externalities by Age Group and Year.
Population $\geq$ 25,000.

	Age Range							Age
	26-30	31-35	36-40	41-45	46-50	51-55	56-60	p-val
1990	0.012^{b}	0.010^{c}	0.010^{c}	0.010	0.005	0.004	0.008	0.000
	(0.006)	(0.006)	(0.006)	(0.006)	(0.006)	(0.007)	(0.006)	
2000	0.019^{a}	0.018^{a}	0.011^{a}	0.007^{c}	0.008^{a}	0.012^{a}	0.009^{a}	0.000
	(0.003)	(0.003)	(0.003)	(0.004)	(0.003)	(0.003)	(0.003)	
2010	0.025^{a}	0.028^{a}	0.027^{a}	0.023^{a}	0.017^{a}	0.016^{a}	0.021^{a}	0.000
	(0.004)	(0.004)	(0.004)	(0.003)	(0.003)	(0.003)	(0.003)	
2019	0.023^{a}	0.023^{a}	0.019^{a}	0.020^{a}	0.018^{a}	0.016^{a}	0.011^{a}	0.000
	(0.003)	(0.002)	(0.003)	(0.003)	(0.003)	(0.004)	(0.004)	
Year p-val	0.037	0.015	0.000	0.000	0.000	0.011	0.023	

Note: ^{a} indicates the coefficient is significant at the 1%, ^{b} at the 5%, and ^{c} at the 10%. All the coefficients come from Table A.13, Column 1. The Table uses the subsample of 804 cities that had a population larger than 25,000 inhabitants in 1980. The last column presents the p-value that comes from testing for each decade if the effect between age groups is constant. The last row presents the p-values that comes from testing if for every age group the effect is constant across decades.

Table A.15: Agglomeration Externalities by Age Group and Year. Population $\geq$ 50,000.

	Age Range							Age
	26-30	31-35	36-40	41-45	46-50	51-55	56-60	p-val
1990	0.012[b]	0.010	0.010	0.009	0.005	0.004	0.007	0.000
	(0.006)	(0.006)	(0.006)	(0.007)	(0.007)	(0.007)	(0.006)	
2000	0.021[a]	0.020[a]	0.012[a]	0.007[c]	0.009[a]	0.012[a]	0.009[a]	0.000
	(0.003)	(0.003)	(0.004)	(0.004)	(0.003)	(0.003)	(0.003)	
2010	0.027[a]	0.031[a]	0.030[a]	0.024[a]	0.019[a]	0.017[a]	0.022[a]	0.000
	(0.004)	(0.004)	(0.004)	(0.003)	(0.004)	(0.003)	(0.003)	
2019	0.023[a]	0.023[a]	0.019[a]	0.019[a]	0.018[a]	0.015[a]	0.009[a]	0.000
	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.004)	(0.003)	
Year p-val	0.202	0.077	0.006	0.000	0.000	0.109	0.039	

Note: [a] indicates the coefficient is significant at the 1%, [b] at the 5%, and [c] at the 10%. All the coefficients come from Table A.13, Column 2. The Table uses the subsample of 530 cities that had a population larger than 50,000 inhabitants in 1980. The last column presents the p-value that comes from testing for each decade if the effect between age groups is constant. The last row presents the p-values that comes from testing if for every age group the effect is constant across decades.

Table A.16: Agglomeration Externalities by Age Group and Year. Population $\geq$ 75,000.

	Age Range							Age
	26-30	31-35	36-40	41-45	46-50	51-55	56-60	p-val
1990	0.014^{b}	0.011^{c}	0.011^{c}	0.011	0.006	0.005	0.009	0.000
	(0.007)	(0.006)	(0.007)	(0.007)	(0.007)	(0.008)	(0.007)	
2000	0.023^{a}	0.021^{a}	0.014^{a}	0.008^{b}	0.011^{a}	0.014^{a}	0.011^{a}	0.000
	(0.003)	(0.003)	(0.004)	(0.004)	(0.003)	(0.004)	(0.003)	
2010	0.029^{a}	0.032^{a}	0.031^{a}	0.026^{a}	0.020^{a}	0.019^{a}	0.024^{a}	0.000
	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.003)	(0.004)	
2019	0.023^{a}	0.023^{a}	0.019^{a}	0.019^{a}	0.018^{a}	0.015^{a}	0.010^{a}	0.000
	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.004)	(0.004)	
Year p-val	0.403	0.150	0.046	0.001	0.007	0.301	0.041	

Note: ^{a} indicates the coefficient is significant at the 1%, ^{b} at the 5%, and ^{c} at the 10%. All the coefficients come from Table A.13, Column 3. The Table uses the subsample of 396 cities that had a population larger than 75,000 inhabitants in 1980. The last column presents the p-value that comes from testing for each decade if the effect between age groups is constant. The last row presents the p-values that comes from testing if for every age group the effect is constant across decades.

Table A.17: Agglomeration Externalities by Age Group and Year.
Population $\geq$ 100,000.

	Age Range							Age
	26-30	31-35	36-40	41-45	46-50	51-55	56-60	p-val
1990	0.014^{b}	0.011^{c}	0.011	0.011	0.006	0.005	0.009	0.000
	(0.007)	(0.007)	(0.007)	(0.008)	(0.008)	(0.008)	(0.007)	
2000	0.024^{a}	0.021^{a}	0.014^{a}	0.008^{c}	0.011^{a}	0.013^{a}	0.011^{a}	0.000
	(0.003)	(0.003)	(0.004)	(0.004)	(0.003)	(0.004)	(0.003)	
2010	0.030^{a}	0.033^{a}	0.032^{a}	0.026^{a}	0.021^{a}	0.019^{a}	0.025^{a}	0.000
	(0.005)	(0.005)	(0.005)	(0.004)	(0.004)	(0.004)	(0.004)	
2019	0.025^{a}	0.024^{a}	0.020^{a}	0.020^{a}	0.019^{a}	0.016^{a}	0.011^{a}	0.000
	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.004)	(0.004)	
Year p-val	0.413	0.183	0.033	0.000	0.004	0.296	0.062	

Note: ^{a} indicates the coefficient is significant at the 1%, ^{b} at the 5%, and ^{c} at the 10%. All the coefficients come from Table A.13, Column 4. The Table uses the subsample of 306 cities that had a population larger than 100,000 inhabitants in 1980. The last column presents the p-value that comes from testing for each decade if the effect between age groups is constant. The last row presents the p-values that comes from testing if for every age group the effect is constant across decades.

A.2.5 Main Results: Different Measure of Density.

In this section I use a different measure of density. Instead of 1980 CBSA population I use (log) 1980 CBSA density.

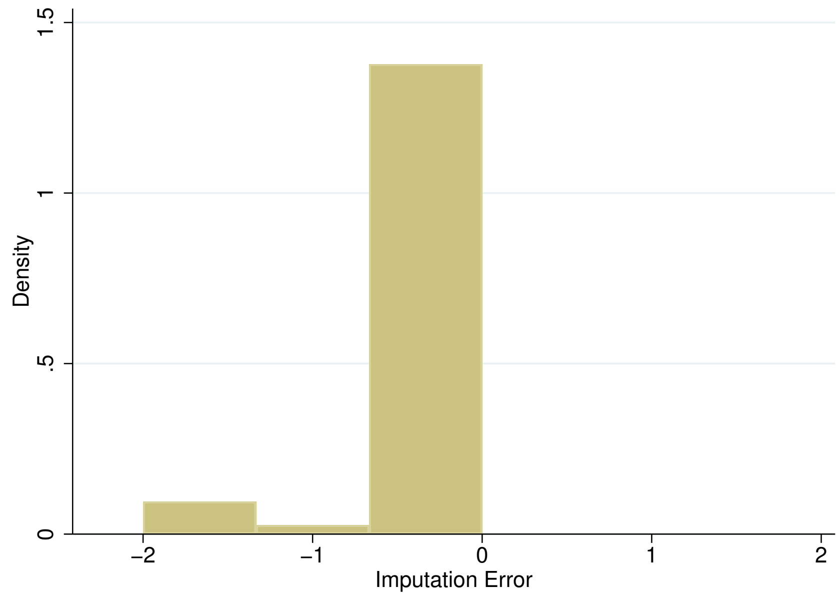
Table A.18: Agglomeration Externalities by Age Group and Year.
Different Measure of Density.

	Age Range							Age
	26-30	31-35	36-40	41-45	46-50	51-55	56-60	p-val
1990	0.029^{b}	0.024^{b}	0.023^{b}	0.023^{b}	0.016	0.015	0.023^{b}	0.000
	(0.012)	(0.012)	(0.011)	(0.011)	(0.011)	(0.011)	(0.010)	
2000	0.037^{a}	0.036^{a}	0.026^{a}	0.019^{b}	0.020^{a}	0.026^{a}	0.021^{a}	0.000
	(0.008)	(0.008)	(0.009)	(0.009)	(0.007)	(0.007)	(0.008)	
2010	0.036^{a}	0.044^{a}	0.046^{a}	0.040^{a}	0.031^{a}	0.029^{a}	0.037^{a}	0.000
	(0.008)	(0.007)	(0.008)	(0.008)	(0.008)	(0.007)	(0.007)	
2019	0.032^{a}	0.033^{a}	0.030^{a}	0.035^{a}	0.034^{a}	0.031^{a}	0.020^{a}	0.000
	(0.007)	(0.006)	(0.006)	(0.006)	(0.006)	(0.007)	(0.007)	
Year p-val	0.187	0.004	0.011	0.014	0.001	0.017	0.054	

Note: ^{a} indicates the coefficient is significant at the 1%, ^{b} at the 5%, and ^{c} at the 10%. The measure of density correspond to 1980 CBSA population divided by CBSA area. The last column presents the p-value that comes from testing for each decade if the effect between age groups is constant. The last row presents the p-values that comes from testing if for every age group the effect is constant across decades.

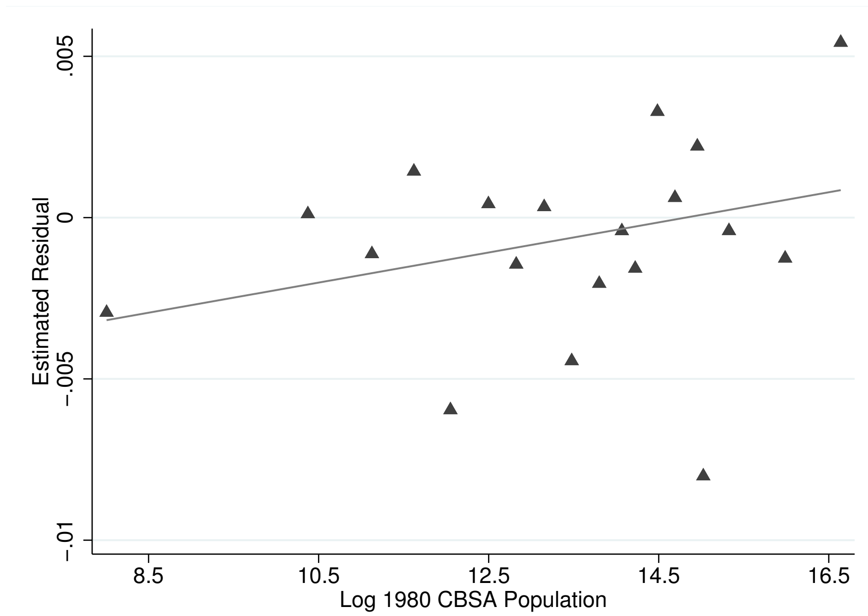
B Appendix Figures

Figure B.1: Weeks Worked Imputation Error



Note: I use the those samples with the exact number of weeks worked (i.e., 1980, 1990, and 2000) to construct the prediction error of imputing 52 as the number of weeks worked.

Figure B.2: Predicted Error Term and Population



Note: Full-time full-year men workers. This graph present a binscatter graph of the error term (multiplied by 100 millions) of a regression of (log) hourly wage on education, city, age, decade, country of origin, race, and industry dummies.

C Model Appendix

The model predicts that the log ratio of college and non-college workers follows:[21]

$$\begin{aligned} log\,(r_{ctj}) = log\left(\frac{\theta_t^s}{\theta_t^u}\right) \quad &+ \quad \left(\frac{1}{\sigma_A} - \frac{1}{\sigma_E}\right) log\left(\frac{S_{ct}}{U_{ct}}\right) \\ &+ \quad (\mu_{tj}^s - \mu_{tj}^u) log\,(D_c) - \left(\frac{1}{\sigma_A}\right) log\left(\frac{S_{ctj}}{U_{ctj}}\right) \end{aligned}$$

[21] This equation correspond to Equation 6 in Section 4.1.

The total derivative of equation 6 follows:

$$
\begin{aligned}
\mathrm{d}log\left(r_{ctj}\right) \quad = \quad & \frac{\partial log\left(r_{ctj}\right)}{\partial log(\theta^s_{ct})}\mathrm{d}log(\theta^s_{ct}) + \frac{\partial log\left(r_{ctj}\right)}{\partial log(\theta^u_{ct})}\mathrm{d}log(\theta^u_{ct}) \\
& + \frac{\partial log\left(r_{ctj}\right)}{\partial log(S_{ct})}\mathrm{d}log(S_{ct}) + \frac{\partial log\left(r_{ctj}\right)}{\partial log(U_{ct})}\mathrm{d}log(U_{ct}) \\
& + \frac{\partial log\left(r_{ctj}\right)}{\partial log(S_{ctj})}\mathrm{d}log(S_{ctj}) + \frac{\partial log\left(r_{ctj}\right)}{\partial log(U_{ctj})}\mathrm{d}log(U_{ctj}) \\
& + \frac{\partial log\left(r_{ctj}\right)}{\partial \mu^s_{tj}}\mathrm{d}\mu^s_{tj} + \frac{\partial log\left(r_{ctj}\right)}{\partial \mu^u_{tj}}\mathrm{d}\mu^u_{tj}
\end{aligned}
\tag{C.1}
$$

Now I will consider each component:

$$\frac{\partial log\left(r_{ctj}\right)}{\partial log(\theta^s_{ct})}\mathrm{d}log(\theta^s_{ct}) = \mathrm{d}log(\theta^s_{ct})$$

$$\frac{\partial log\left(r_{ctj}\right)}{\partial log(\theta^u_{ct})}\mathrm{d}log(\theta^u_{ct}) = -\mathrm{d}log(\theta^u_{ct})$$

$$\frac{\partial log\left(r_{ctj}\right)}{\partial log(S_{ct})}\mathrm{d}log(S_{ct}) = \left(\frac{1}{\sigma_A} - \frac{1}{\sigma_E}\right)\mathrm{d}log(S_{ct})$$

$$\frac{\partial log\left(r_{ctj}\right)}{\partial log(U_{ct})}\mathrm{d}log(U_{ct}) = -\left(\frac{1}{\sigma_A} - \frac{1}{\sigma_E}\right)\mathrm{d}log(U_{ct})$$

$$\frac{\partial log\left(r_{ctj}\right)}{\partial log(S_{ctj})}\mathrm{d}log(S_{ctj}) = -\frac{1}{\sigma_A}\mathrm{d}log(S_{ctj})$$

$$\frac{\partial log\left(r_{ctj}\right)}{\partial log(U_{ctj})}\mathrm{d}log(U_{ctj}) = \frac{1}{\sigma_A}\mathrm{d}log(U_{ctj})$$

$$\frac{\partial log\left(r_{ctj}\right)}{\partial \mu_{tj}^{s}}\mathrm{d}\mu_{tj}^{s} = log(D_c)\mathrm{d}\mu_{tj}^{s}$$

$$\frac{\partial log\left(r_{ctj}\right)}{\partial \mu_{tj}^{u}}\mathrm{d}\mu_{tj}^{u} = -log(D_c)\mathrm{d}\mu_{tj}^{u}$$

Plugging these equation into Equation C.1 and grouping similar terms we arrive to:

$$\begin{aligned}\mathrm{d}log(r_{cj}) = \mathrm{d}log\left(\frac{\theta^s}{\theta^u}\right) - \frac{1}{\sigma_E}\mathrm{d}log\left(\frac{S_c}{U_c}\right) + (\mathrm{d}\mu_j^s - \mathrm{d}\mu_j^u)log\left(D_c\right) \\ + \left(\frac{1}{\sigma_A}\right)\left(\mathrm{d}log\left(\frac{S_c}{U_c}\right) - \mathrm{d}log\left(\frac{S_{cj}}{U_{cj}}\right)\right)\end{aligned} \tag{C.2}$$

which is Equation 7 in the text.