Workplace Agglomeration and Social Network Segregation:

Labor Market Returns by Race

Elizabeth Ananat

Duke University

Shihe Fu

Xiamen University

Stephen L. Ross

University of Connecticut

December 14, 2011

Abstract

In this paper, we document that wages of nonwhites, and particularly of blacks, appear to rise less with agglomeration of employment and concentrations of human capital than do white wages. For blacks, this pattern holds even though our method allows for non-parametric controls for the effects of age, education, and other demographics on wages and for the return to agglomeration and human capital concentrations to vary across the same demographic variables and across metropolitan areas. We find that an individual's return to agglomeration in wages rises with the share of workers in a work location who have the same race as this individual. This finding is consistent with non-whites receiving lower returns to agglomeration and human capital concentrations because they have fewer same-race peers and fewer highly-educated same-race peers at work from whom to enjoy spillovers. As further support for this hypothesis, we estimate models of total factor productivity using data on manufacturing establishments for the same sample of metropolitan areas. We find evidence that the relationship between firm productivity and agglomeration increases in magnitude when the race composition of the firm's employees matches the race composition of other workers in the same location.

I. Introduction

Two of the most salient characteristics of American cities are agglomeration economies—cities exhibit higher productivity (Ciccone and Hall 1996; Henderson 2003) and wages (Glaeser and Mare 2001) than do less-urbanized areas—and high rates of segregation and inequality, particularly around race, with African-Americans facing high levels of racial segregation and earning substantially less than do whites (Ananat 2011). An open question in the literature is whether one component of these racial pay disparities is that minorities and whites derive different benefits from agglomeration, and if so why.

Using most datasets, it is difficult to identify the effect of work location on wages and therefore also difficult to examine whether and when racial differences exist in the return to agglomeration and other location attributes, such as human capital externalities. Individuals choose, at least to some extent, where in a city to live and to work, so residential and workplace segregation and networks, as well as distance between residence and workplace, are endogenous. Moreover, individuals may select into particular neighborhoods based on skill and human capital, and many important aspects of these individual characteristics are unobservable to the econometrician.

Work by Fu and Ross (2010), however, exploits the fact that households systematically sort over residential location and demonstrates that residential location fixed effects provide an effective control for unobserved ability using restricted Census data that includes individual workplace and residential location. Specifically, Fu and Ross (2010) find little evidence of bias in their estimates of within-metropolitan-area agglomeration effects from sorting across workplaces, and consistent with this conclusion they also demonstrate that the within-metropolitan-area correlation between observable ability and agglomeration is very low. Further, while returns to human capital externalities are attenuated by the inclusion of residential fixed effects, additional analyses confirm that the remaining estimated effects of human capital externalities are unlikely to be driven by unobserved ability because observationally equivalent workers in different work locations are earning similar wages net of commuting costs. Finally, the use of residential fixed effects reduces unexplained racial differences in wages by 53%, which is comparable to the 48% reduction in the race coefficient found by Lang and Manove (2006) from the inclusion of the AFQT score as a measure of ability.

In this paper, we document that wages of nonwhites, and particularly of blacks, appear to rise less with agglomeration than do white wages for a sample of prime age, fully employed males residing in metropolitan areas with more than one million residents. For African-Americans, this pattern holds even though our method allows for non-parametric controls for the effect of observables such as age, education, and other demographics and unobservables through residential location on wages, and allows for the return to agglomeration to vary across the same demographic variables, industry, occupation, and metropolitan areas. Further, these racial differences cannot be explained by any of the individual choices that we explore, such as whether to use mass transit, the racial composition of their residential neighborhood, or the density of the location in which they work. Notably, the racial differences in returns are significantly larger in industries that have the largest returns to agglomeration or to human capital externalities.

Next, we explore whether these differences in returns might be explained by race-specific information networks (Hellerstein et al. 2009; Ionnides and Loury 2004). We find that higher own-race representation in a work location and among the college-educated workforce in that location increases the returns to agglomeration. These results are consistent with non-whites receiving lower returns to agglomeration because they have fewer same-race peers and fewer highly-educated same-race peers at work from whom to enjoy spillovers and so are relatively less productive in such workplaces. Notably, both the black-white differences in returns to agglomeration and in human capital spillovers are zero after controlling for differences in returns based on own race representation in the work location.

Finally, in order to provide additional evidence that racial networks affect productivity spillovers, we estimate total factor productivity (TFP) models for manufacturing establishments covering the same metropolitan areas as our worker sample. Following Moretti (2004), we identify a sample of workers in each establishment based on zip code-three digit industry cells, which allows us to estimate a trans-log production function that includes controls for the education level of an establishment's workplace. As Moretti found for human capital externalities when examining metropolitan-area level variation, we find that firm TFP increases in locations that have high concentrations of employment and human capital. Further, we find that the productivity returns to agglomeration and to human capital externalities fall substantially (only the agglomeration difference is statistically significant) when the race of the firm's workers does not closely match the racial composition of the surrounding location.

The rest of the paper proceeds as follows. Section II reviews the literatures on agglomeration economies and on the causes of wage disparities. Section III describes our wage model. Section IV describes the individual data, and section V presents our wage model results. Section VI presents our TFP models, and section VII discusses and concludes.

II. Literature review

A large and diverse literature documents the disadvantages and adverse outcomes experienced by African-Americans in segregated neighborhoods and metropolitan areas.

African-Americans face much higher levels of residential segregation and centralization than other minority groups (Massey and Denton 1993), and adverse changes in U.S. central cities over the last several decades may have disproportionately affected African-Americans. Wilson (1987) argues that African-Americans' outcomes are negatively affected by their concentration in increasingly poor and distressed central city neighborhoods. Kain (1968) suggests that the increasingly poor job access of African-Americans in central cities may have adverse effects, and recent work by Hellerstein et al. (2008) finds that the influence of employment access is race-specific, so that employment depends heavily on access to places where members of one's own race are employed.[1] Further, Cutler and Glaeser (1997), Card and Rothstein (2007) and Ananat (2011) provide evidence that segregation leads to worse education and labor market outcomes for African-Americans, and similarly Edin et al. (2003) and Damm (2006) find that the placement of refugees into ethnic enclaves in Sweden and Denmark, respectively, affects labor market outcomes.

Beyond the potential social influences of location, labor market outcomes are directly influenced by relationships between workers. Suggesting the importance of referrals, Bayer et al. (2008) find that similar individuals who reside on the same block are more likely to work together and that similarity of a worker to others residing nearby influences both employment and wages. Hellerstein et al. (2009) find that employees at the same firm are more likely to come from the same neighborhood than are employees who work at different firms in the same location. This effect primarily operates within racial and ethnic groups. Dustman et al. (2009) find that minority workers in Germany are much more likely to work in locations where other minorities work. They find that workers benefiting from referrals earn higher initial wages due to

[1] The literature associated with the spatial mismatch hypothesis is huge. See Ihlanfeldt and Sjoquist (1998) and Kain (1992) for detailed surveys.

the information revealed by the referral but have slower wage growth, presumably since firms learn less about workers over time.[2]

Further, substantial evidence documents peer effects in the workplace. Peers can affect one another's productivity through establishing norms about absenteeism (Bokenblom and Ekblod 2007; Ichino and Maggi 2000; Lindbeck et al. 2007; DePaola 2008) or work effort (Bandiera, Barankay and Rasul 2005; Falk and Ichino (2006); Mas and Moretti 2006). Nanda and Sorenson (2008) find evidence of peer effects on self-employment suggesting knowledge or experience sharing between workers.[3]

To the extent that agglomeration economies arise from spillovers across individuals, it stands to reason that the productivity and wage benefits of such spillovers could depend upon the quality of the social interactions available around and near a workplace. Focusing first on firms, Duranton and Puga (2001) suggest that firms learn about production processes from being located in diverse cities, and Rosenthal and Strange (2003) find that the likelihood of firm births is increased by the geographic proximity of other firms in the same industry, especially within the first mile, suggesting a substantial role for social interactions. Ellison, Glaeser and Kerr (2010) find evidence that spillovers between firms explain a significant portion of the co-agglomeration of industries using metrics for the extent that firms share workers and ideas. Audretsch and Feldman (1996) and Feldman and Audretsch (1999) demonstrate that the composition of surrounding industry affects the rate of product innovation. Finally, Moretti

[2] Ioannides and Loury (2004) provide a detailed review of the extensive literature on labor market referrals and networks documenting several important stylized facts. Also see Granovetter (1995).
[3] See Ross (In Press) for a recent review of the general literature on neighborhood and peer effects.

(2004) finds that firms are more productive and more innovative when located in cities that have more educated workers even after controlling for the education level of the firm's workforce.[4]

The most direct evidence of workers' productivity being influenced by surrounding firms, workers and/or economic activity arises from models that examine the influence of agglomeration on wages. Glaeser and Maré (2001), Wheeler (2001), Combes, Duranton, and Gobillon (2004), Rosenthal and Strange (2006), Yankow (2006), Fu (2007), Di Addario and Patacchini (2008), and Fu and Ross (2010) all find that wages are higher in large labor markets with high concentrations of employment, and many of these studies also find a positive link between wages and the human capital level associated with an employment concentration.[5] Notably, Glaeser and Maré (2001) find that workers who migrate away from large metropolitan areas retain their earnings gains and argue that these permanent gains arise because workers learn more skills when working in dense urban areas. Rosenthal and Strange (2006) and Fu and Ross (2010) find evidence that employment density and human capital spillovers arise within metropolitan areas. Rosenthal and Strange document a fairly rapid decay of these spillovers across space, again consistent with agglomeration resulting from social interactions, as opposed to deriving from shared infrastructure or externalities associated with a broader labor market.

Human capital spillovers at work can stem from a variety of peer and social interaction effects that may not be that different from the types of interactions that arise within a firm. As discussed above, peers can affect one another's productivity through establishing norms about absenteeism or work effort. Also, social peers may influence productivity via the sharing of knowledge (Nanda and Sorenson 2008). If peers share knowledge not only about how to be

[4] See Audretsch and Feldman (2004), Duranton and Puga (2004), Moretti (2004) and Rosenthal and Strange (2004) for detailed surveys of the literature on agglomeration economies and production externalities within cities.
[5] Other studies, including Wheaton and Lewis (2002), Fu (2007), and Combes, Duranton, and Gobillon (2004) find evidence that wages increase with concentrations of employment in an individual's own occupation or industry.

productive on the job, but also about job opportunities, match quality may be greater in denser areas. To the extent that, even within the same industry, individuals are more likely to associate with peers of the same race, race-specific knowledge networks (Hellerstein et al., 2009) and job-finding networks could explain why in most industries (where whites make up the bulk of workers), human capital spillovers may accrue more to whites than to nonwhites.

Moreover, an individual's residential environment can affect his ability to take advantage of workplace agglomeration. Massey and Denton (1993) hypothesized that racial segregation, combined with lower average earnings among blacks than among whites, reduces both norms that promote career success (Lindbeck, Palme, and Persson 2007) and access to jobs among African-Americans. Further, spatial mismatch, i.e. the concentration of African-Americans in central cities in combination with decentralization of employment, is thought to reduce blacks' access to jobs and limit their labor market opportunities (Kain, 1969). One mechanism by which job network access might influence employment, wages, and worker productivity is by its influence on the racial distribution of employment over space. For example, Hellerstein, Neumark, and McInerney (2008) find that an African-American's chance of getting a job is most influenced by his access to firms where other African-Americans currently work.

In this paper, we test whether racial disparities exist in the return to agglomeration. We also test whether own-race share of employment in the area where an individual works moderates the racial disparity in return to agglomeration. As part of these tests, we examine two types of agglomeration economies. The first, captured by the density of employment in the part of an MSA in which an individual works (the "workplace PUMA"), focuses on general spillovers associated with the total amount of economic activity in an area. The second, captured by the share of workers in an individual's workplace PUMA who are college graduates, focuses on skill-based human capital spillovers.

III. Model

First, to establish a baseline measure of agglomeration economies, we estimate the equation:

(1) $y_{ijk} = \beta_j Z_k + \alpha_j X_i + \pi_i + \varepsilon_{ijk}$

to predict the log wage y_i of worker i working in area k. Z_k, our measure of local agglomeration economies, represents either density of employment in the area where worker i is employed, or share of workers with a college degree in the area where worker i is employed. X_i is a vector of worker characteristics, including race and ethnicity, age, education, family status, nativity, share of workers in worker i's industry in the MSA who have a college degree, and share in worker i's occupation in the MSA who have a college degree. π_i is a fixed effect for worker i's metropolitan area. In some specifications, we remove controls for race and instead estimate equation (1) separately by race and ethnicity, in order to demonstrate the average differences in returns to agglomeration by race.

Second, our main analysis collapses the individual data to observationally equivalent cells, indexed by j, based on age, race, family status, education, and nativity plus the residential census tract:

(2) $y_{ijk} = \beta_j Z_k + \delta_j + \varepsilon_{ijk}$

where the inclusion of cell fixed effects, δ_j, mitigates bias from unobservables because observationally equivalent workers who choose the same residential location should be expected to have similar unobservables as well (Fu and Ross, 2010). By mean differencing, we get

(3) $y_{ijk} - \bar{y}_j = \beta_j (Z_k - \bar{Z}_j) + (\varepsilon_{ijk} - \bar{\varepsilon}_j).$

Note that we allow the effect of agglomeration, β_j, to be heterogeneous within the sample. (We have estimated this model with a constant β and find robust, precisely estimated agglomeration effects.)

We can then create:

(4) $\hat{\beta}_{ij} = \frac{y_{ijk}-\bar{y}_j}{Z_k-\bar{Z}_j} = \beta_j + \frac{\varepsilon_{ijk}-\bar{\varepsilon}_j}{Z_k-\bar{Z}_j}$,

where this ratio contains β_j plus an error that is not a function of individual attributes. Therefore, we can parameterize β_j as a function of observables including metropolitan fixed effects (π_j) and estimate the equation

(5) $\hat{\beta}_{ij} = \alpha X_{ij} + \rho_j + \varepsilon_{ij}$

which captures the effect of worker i's characteristics on worker i's return to agglomeration in wages when working in area j. This transformation allows us to estimate a model in which the return to agglomeration is quite heterogeneous including allowing variation in returns across metropolitan areas ρ_j without the use of large numbers of interaction terms. However, the transformation also leads to a very heteroskedastic disturbance term and so our models are estimated using GLS where the variance of the disturbance is estimated for each cell based on the scale factor $(Z_k - \bar{Z}_j)$.[6] In some specifications, we estimate equation (5) separately by certain worker and MSA characteristics.

Naturally, when examining the effect of agglomeration economies, our model should control for human capital externalities and vice versa. In a general model with two variables, Z_{1k} and Z_{2k}, capturing work location spillovers, the model can be written as

(6) $\hat{\beta}_{ij} = \frac{y_{ijk}-\bar{y}_j}{Z_{1k}-\bar{Z}_{1j}} = \beta_{1j} + \rho_j + \beta_{2j}\frac{Z_{2k}-\bar{Z}_{2j}}{Z_{1k}-\bar{Z}_{1j}} + \frac{\varepsilon_{ijk}-\bar{\varepsilon}_j}{Z_k-\bar{Z}_j}$

---

[6] We also verify that estimates of equation (5) are very similar to estimates arising from the inclusion of interaction terms in equation (3) for models that omit the metropolitan fixed effects and so have fewer interaction terms.

where the estimates of the components of β_{2j} are obtained by interacting the linear demographic controls with the ratio of the two mean differenced agglomeration variables. In practice, we estimate models both with the agglomeration variable defined as Z_1 and with the human capital externality variable defined as Z_2, always reporting the coefficient estimates associated with Z_1 in equation (6).[7]

Finally, we estimate:

$$(7) \quad \hat{\beta}_{ij} = \alpha X_{ij} + \rho_j + \theta OR_{ij} + \varepsilon_{ij}$$

where OR_{ij} is the share of workers employed in location j who belong to the same race or ethnicity as individual i (or share of college educated workers in our model of the return to human capital externalities). These estimates allow us to test whether workers get a larger return to agglomeration when more of the workers in the area (or more of the skilled workers in the area) are of the worker's own race.

IV. Data

The models in this paper are estimated using the confidential data from the Long Form of the 2000 U.S. Decennial Census. The sample provides detailed geographic information on individual residential and work location. A subsample of prime-age (30-59 years of age), full time (usual hours worked per week 35 or greater), male workers is drawn for the 49 Consolidated Metropolitan and Metropolitan Statistical Areas that have one million or more residents.[8] These restrictions lead to a sample of 2,343,092 workers, including 1,705,058 whites, 226,173 blacks, 264,880 Hispanics, and 135,577 Asians.

---

[7] The fixed effects allow for variation in the return to Z_1 by metropolitan area and so we prefer our estimates of parameters based on Z_1, but estimates when the same variable is assigned to Z_1 are similar.

[8] This sample is comparable to the sample drawn from the Public Use Microdata Sample (PUMS) of the 2000 Census by Rosenthal and Strange (2006) except that we explicitly restrict ourselves to considering residents of mid-sized and large metropolitan areas.

Table 1 reports individual, employment location PUMA [9] characteristics, and metropolitan area characteristics by race[10] (white, African-American, Hispanic, or Asian) of the worker. Our dependent variable is the logarithm of the wage, which is based on an individual's labor earnings last year divided by the product of the number of weeks worked and the average hours per week worked last year. Our demographics controls include categorical variables by age, education, family structure, and immigration status. These controls are also used to create the observationally equivalent cells described above. At the employment PUMA, we measure agglomeration with employment density and human capital externalities by calculating the share of workers with at least four years of college education based on all full time workers reporting this employment location. These two variables are created based on both overall employment and employment in a worker's one digit industry code. The share of own race variables are also constructed using all full time workers.

We also generate additional worker characteristics that are not included among our demographic controls, but are used to split the sample for supplementary analyses. The variables include mass transit versus car commuters; central city versus suburban residents; those living in census tracts with high versus low shares African-American; those working in highly dense versus less dense areas of their MSA; those working in areas with high versus low concentrations of college graduates; those working in the central city versus the suburbs; and those working in high versus low agglomeration or human capital spillover industries. The last of these measures

[9] PUMA is the Public Use Microdata Area defined to report residential location in the Public Use Microdata Sample (PUMS) of the 2000 Decennial Census and is constructed to contain a population of 100,000 residents. We calculate workplace PUMA variables based on the population of workers reporting their work location at the census tract level, which is then matched to 2000 PUMA definitions. The PUMS also reports individual workplace using an alternative workplace PUMA, but these definitions vary dramatically across metropolitan area. Fu and Ross (2010) confirm that agglomeration estimates are robust to alternative workplace definitions including workplace PUMA, PUMA, and zip code area.

[10] Throughout the paper we use the term "race" interchangeably with "race and ethnicity" to capture distinctions between non-Hispanic whites ("whites"), non-Hispanic blacks ("blacks" or "African-Americans"), non-Hispanic Asian-Americans ("Asians"), and Hispanics of any race ("Hispanics").

was created by estimating the return to either employment density or share college-educated workers for the whole sample in the externality models from equation (5), controlling for a set of industry dummy variables.

V. Results

Table 2 reports the results of basic agglomeration economies models using either overall employment, in column 1, or industry-specific employment, in column 2. The regressions control for education levels in the MSA overall, as well as a variety of individual characteristics (age, race, education, family structure, and nativity). As expected, both employment density and share of workers with a college degree in an individual's place of employment strongly predict higher wages for that individual in both models, confirming the existence of agglomeration economies. The standardized effects for employment density are quite similar, with 0.020 for overall density and 0.025 for industry-specific density, which is not surprising given that the two density variables have a correlation of 0.82. The correlation between overall share college educated and industry share college educated, however, is only 0.62, and the standardized effect of industry-specific share college educated is 0.066, almost double the standardized effect of 0.035 for the overall measure. In addition, we estimate a model including all four agglomeration variables, finding that industry-specific employment density again has a somewhat larger standardized estimate than overall employment density and that industry-specific share college educated captures all the effect of share college educated on wages. Standard errors are cluster at the level of the employment location PUMA.

As shown in Table 3, in which the model is estimated separately for individual racial and ethnic groups, non-whites also receive less of a wage premium for agglomeration, as captured by either employment density or share college educated. Compared to Table 2, which estimates the returns to density and share of workers with a college degree for all workers, Table 3 estimates these returns separately by race, and reveals that whites receive higher than average returns, while nonwhites receive much lower than average returns. African-Americans, in particular, get returns only between about one-fourth and one-third as large as the white return for either the model based on overall employment or the model based on industry-specific employment. For the industry specific workplace variables, a one standard deviation in agglomeration is associated with an increase the black-white wage gap of 1.9 percentage points and one standard deviation share college is associated with a 5.0 percentage point increase. These effects are quite sizable in magnitude when compared to the 14.6 percentage point wage gap in Table 2, and very large when compared to the 6.9 percentage point gap using our preferred wage model shown in Table 5. From this point forward, our analyses will focus on the more empirically relevant industry-specific employment variables, but similar results are found using overall employment.

In Table 4, we investigate the returns to employment density and share college educated as a function of all worker demographics plus industry, occupation, and MSA fixed effects using the model specified in equation (6). This model controls non-parametrically for the effect of observables and unobservables on wages using census tract by demographic cells (plus industry and occupation fixed effects), and allows the return to agglomeration and human capital externalities to vary by education, family structure and immigration status, factors that might explain the differential return to agglomeration across racial and ethnic groups. The dependent variables in this table represent the estimated change in wages for a unit change in density or share college, i.e. the slope between the log wage and each variable, from equation (5). Coefficients in this table report the change in these slopes with a change in worker demographics, answering the question: how does the return to agglomeration differ for workers with different characteristics? Standard errors are cluster at the census tract level.

The table reveals that the return to employment density differs little by age, education, or immigration status; while these are significant drivers of wages themselves, they do not appear in most cases to greatly affect the relationship between wages and employment density (with the exception of obtaining a degree beyond a masters degree). For share college educated, returns to agglomeration rise with both age and education. Turning to the race and ethnicity variables, blacks continue to receive a substantially lower return to either employment density or share college educated in the place of employment. The estimated racial differences in return to both employment density and share college are greater than the effect of having obtained a degree beyond a masters degree relative to being a high school graduate. In comparison, in the wage model from Table 2, the effect of education beyond a master's degree on wages is more than three times the racial differences in wages. These estimated differences in return to agglomeration between blacks and whites cannot be explained by simple differences in unobserved productivity. On the other hand, Hispanics and Asians, after addressing unobservables, do not earn significantly lower wage premiums than whites.[11]

Next, we explore concerns about unobservable ability differences and our model structure in more depth. In Table 5 panel 1, we present the race coefficients from a wage model controlling for a variety of fixed effect specifications: census tracts, block groups, tract by demographic (excluding race) cells, tract by demographics by industry cells, tract by demographics by whether the industry has a high or low return to density, tract by demographics by whether the industry has a high or low return to employment density and/or share college by

[11] Note that the R-squared's are very small in the slope-model regressions. This is a feature of the slope models and is consistent with small changes in the R-squared's when interaction terms are added to wage models.

blue or white collar occupation. The race coefficient for blacks falls by 53%, from -0.15 to -0.07, with the inclusion of tract fixed effects. By comparison, Lang and Manove (2006) include AFQT score as a measure of ability and find that the inclusion of this control lowers the race coefficient on wages by 48% from -0.29 to -0.15. The race coefficient is then fairly stable across all other specifications. In other words, the additional controls do little to explain the remaining racial differences in wages. For Hispanics, however, the inclusion of tract by demographic fixed effects does significantly reduce the ethnic differences in wages.

Panels 2 and 3 show the return to density and the return to share college educated. Both estimates for blacks are quite stable as either geography is refined to the block group level or as tract by demographic cell fixed effects are included. The racial differences in the return to college educated are robust to more detailed cell structures based on industry and occupation, but the racial differences in employment density are not robust to the use of industry and occupation in the cell structure. However, we believe that it is significant that these alternative cell structures lead to dramatic decline in the relevant sample. The tract by demographic cell structure results in 37 percent of the sample residing in singleton cells, which do not contribute to our estimates, while the cell structures involving industry fixed effects or industry and occupation lead to 76 and 60 percent of the sample residing in singleton cells. Given that the enhanced cell structure does little to control for racial differences in wages, it is reasonable to question whether the changes are due to a more detailed specification or instead to the sample selection that results from this more detailed specification.

In Table 6 Panel 1, we present estimates for wage, density and share college models as a function of the standard linear controls including demographics, industry, occupation, and census tract plus a dummy variable for whether an observation belongs to a singleton cell based on the three cell structures from Table 5. While the singleton dummy variable is significant in all wage and employment density models, the coefficient estimates for the tract by demographic cell structure is significantly smaller for all three models than the estimates for the other two cell structures. In panel 2, we estimate second stage employment density and share college slope models based on the linear first stage models presented in panel 1, including a dummy variable for being in a singleton cell based on a given cell structure. The singleton cell estimates for the tract-demographic cell structure in column 1 are small and less statistically significant than the estimates for the industry-based cell structures in columns 2 and 3. For the rest of our analysis, we proceed with the tract by demographic cell model specification, while recognizing that our density results on race are not quite as robust as our share college educated results.

Even if our estimated racial differences cannot be explained by unobserved productivity variables, they might be affected by racial differences in tastes that cause whites and blacks to make different choices that in turn might influence the return to agglomeration. Table 7 reports our estimates of the racial and ethnic differences in the return to agglomeration for several subsamples that reflect variables/choices that are not part of our empirical specification. For example, we know that African-Americans are more likely to use mass-transit than whites, so if mass transit riders were exposed to different employment environments on average and as a result experienced lower returns to agglomeration, then the racial differences might not persist when the sample is split by mass transit usage. We estimate our models in Table 4 for subsamples based on the following criteria: mass transit vs. automobile commuters, central city vs. suburban workplace, residential tract share black above/below metropolitan average,[12] central

[12] We also examine a sample split based on residing in a racially segregated metropolitan area, rather than a high share black neighborhood. However, racial segregation is highly correlated with share black and the majority of our black sample resides in areas that are above the median in terms of segregation. Our estimates for the above median segregation level metropolitan areas are very similar to our results in Table 4, but our estimates for the below

city vs. suburban residence, workplace employment density above/below metropolitan average, workplace share college above/below metropolitan average, worker in a high/low return to employment density industry, and worker in a high/low return to share college industry.[13]

The first noteworthy result from Table 7 is that for each pair of subsamples the effect of employment density and the effect of share college educated is always statistically significant for at least one of those subsamples. Therefore, the estimated differences in returns cannot be simply attributed to minorities selecting into a population that is exposed to lower returns to agglomeration. The racial differences in return to share college are statistically significant or very close to significance at the 5% level for all subsamples. While the racial differences in return to employment density are zero or even positive in some subsamples, the differences in the race coefficient for employment density between the subsamples are never statistically significant except for the split by worker in a high/low return to employment density industry.

Further, the results for the industry splits are supportive of our conclusion that blacks are getting a lower return to employment density and share college educated in work location. In the return to employment density split, blacks in both samples get a lower return to employment density, but the racial difference is significantly larger (t-statistics of 1.95) in the industries where the return to density is larger. Similarly, in the return to share college educated split, the racial difference in larger (t-statistic of 2.89) in the industries where human capital spillovers are larger. These findings are consistent with blacks getting a lower return to the productivity

---

median subsample are much noisy and less stable. For example, the race estimate on agglomeration in the low segregation subsample changes dramatically depending upon whether we use 2000 or 1970 measures of segregation.

[13] Three of our 15 one digit industries (construction industry; arts, entertainment, recreation and food; personal and repair services) are classified as high returns from employment density. The fixed effect estimates are 0.0096, 0.0192 and 0.0071, respectively, above the median estimate and the next closest industry was only 0.0028 above the median estimate. Three industries (transportation and warehousing; finance, insurance and real estate; personal and repair services) are also classified as high returns from share college educated. The fixed effect estimates are 0.214, 0.197, and 0.128, respectively, above the median estimate and the next closest industry was only 0.069 above the median estimate. Most fixed effect estimates that were below the median were also relatively close to the median.

benefits of employment density and share college educated. Note that these splits are based on the estimates of the industry fixed effects for the models presented in Table 4.

Considering the Racial Composition of the Workforce

Table 8 tests whether the pattern of racial wage disparities is consistent with work location spillovers that arise as part of race-specific networks. Under such circumstances, nonwhites may be disadvantaged because they lack same-race peers in the area where they work. In order to examine this hypothesis, we control for own-race share to test whether it moderates racial differences in the return to agglomeration economies. In models of the return to overall employment density, we control for own-race share of employment in the PUMA of employment; in models of return to share college educated, we control for own-race share of college educated workers in the PUMA of employment. Not surprisingly, the main effect of each of these controls is positive and highly significant, consistent with own-race workplace networks increasing the return to agglomeration.

As with the estimated racial differences, these effects are sizable in magnitude. multiplying the coefficient estimate on share own race in workplace times the black-white gap in share own race yields an effect of 0.008, which is almost as large as the 0.009 raw black-white difference in the return to agglomeration. Similarly, multiplying the coefficient estimate on share own race with college in workplace times the black-white gap in share own race college yields an effect of 0.097, which is about 1/3 of the 0.294 racial difference in return to share college.[14] Scaling these estimates shows that one standard deviation changes in employment density and share college are expected to increase the black wage gap by 1.9 and 1.6 percentage points respectively.

[14] The smaller share of the raw effect arises primarily because while the tract by cell fixed effects has little impact on the estimated return to employment density, these controls substantially erode the return to share college.

Further, including own race share in a regression explains almost the entire racial differences in the return to employment density and share college educated. We find that for overall employment density, controlling for own-race share reduces racial differences in return from -0.0061 to -0.0002, and for share college educated controlling for own-race share reverses the racial differences from -0.1678 to 0.0624. Both black coefficients are effectively zero in the models that control for own race share. These findings complement earlier findings by Hellerstein et al. (2009) that employment networks operate along racial lines. For Hispanic and Asian workers, the estimates suggest that they are in fact overcompensated relative to whites, although only the differences for share college educated are significant at the 0.05 level.

Table 9 shows the specification from Table 8 run separately for workers in industries that have high/low returns to employment density or for workers in industries that have high/low returns to share college educated. Although the differences are not statistically significant, the qualitative pattern is exactly as one would expect: the effect of own share in work location on return to employment density is substantially larger in industries with high returns to employment density (0.0277 vs. 0.0136) and the effect of own share college educated workers in work location on return to share college educated is substantially larger in industries with high return to share college educated (0.5617 vs. 0.3594).

Do Racial Networks affect Productivity

In order to examine whether racial networks affect productivity, we turn to estimating models of firm productivity using establishment data gathered in the 1997 census of manufacturers. We restrict ourselves to manufacturing data because information on the cost of materials and on the stock of capital, which is necessary to estimate productivity, is only available for the manufacturing industry. Prior to estimating models of firm productivity, we first need to verify that the pattern of racial differences observed in our sample of all workers also arises in a sample of manufacturing workers.

Table 10 presents the estimates for the models presented in Table 8 for the subsample of manufacturing workers. Our estimates are not sufficiently precise to identify statistically significant racial differences, but the magnitudes of our estimates are comparable, with racial differences of -0.0129 and -0.1008 for employment density and share college in the manufacturing subsample as compared to -0.0061 and -0.1678 for the full sample. Similarly, our estimates on own race share in the manufacturing sample are 0.0201 and 0.2510 for employment density and share college versus 0.0151 and 0.3756 for the full sample. The effect sizes are larger for employment density and smaller, but still between 60 and 70 percent of the magnitude, for share college educated, which is not surprising because the return to share college educated in manufacturing was smaller than many other industries.[15]

In order to estimate a model of total factor productivity, we access the population of establishments in the same sample of metropolitan areas with populations over one million using the 1997 Census of Manufacturers. Using these data, we can estimate models for firm net revenues (minus material costs) as either a Cobb-Douglas or a translog function of structure capital, equipment capital, and employment. For employment, we follow Moretti (2004) and Hellerstein, Neumark, and Troske (1999) and develop estimates of the share of workers at a firm with four year college degrees based on the decennial Census. This share is combined with firm total employment to estimate the number of college and non-college educated workers. Because our analysis looks within metropolitan areas and we have confidential data for both the decennial Census and the Census of Manufacturers, we are able to estimate the number of college and non-

[15] While manufacturing is not in the high return category for either density or share college, the density return fixed effect estimate for manufacturing is above the median fixed effect estimate by 0.0028, while the share college return fixed effect estimate is below the median fixed effect estimate by 0.042.

college workers in each firm based on placing firms into three digit industry code by zip code cells, rather than industry code by metropolitan area cells as done by Moretti (2004). When we cannot match establishment zip code to decennial Census data, we base our estimates on industry-PUMA cells. All models control for three digit industry and metropolitan area fixed effects, and standard errors are clustered at the level of the work location PUMA.

Table 11 Columns 1 and 2 show the results for models similar to Moretti's using Cobb-Douglas and Translog production functions, respectively.[16] We estimate that the effect of a one standard deviation increase in the share college educated workers in a PUMA on firm total factor productivity is 0.020, which is comparable in magnitude to Moretti's cross-sectional estimates that a one standard deviation increase in share college educated increases total factor productivity by between 0.035 and 0.049—especially when considering that our estimate is reduced by our inclusion of a control for employment density. The Translog fits the data substantially better than Cobb-Douglas with the R-Squared increasing from 0.84 to 0.91, which is a huge increase given the relatively small change in available degrees of freedom. The resulting F-statistic is 8,147 dramatically rejecting the Cobb-Douglas model. While the estimate on the share college is relatively unchanged, the employment density estimate increases by 0.0011, which is sizable given that the standard errors on the estimates fall between 0.0001 and 0.0002. Further, the translog model yields much more precise estimates on both employment density and share college with the standard errors fall by 30 and 35 percent, respectively.[17]

[16] Our estimates on log capital equipment, log capital structure, log unskilled labor and log skilled labor are 0.526, 0.036, 0.293 and 0.048 as compared to Moretti's estimates of 0.178, 0.470 and 0.322 for 1992 and 0.476, 0.333 and 0.196 for 1982 for log capital, log unskilled labor and log skilled labor. Our within metropolitan area estimates for capital equipment and unskilled labor are quite similar to Moretti's estimates for 1982. While we replicate Moretti's finding of lower return to skilled labor, our coefficient on skilled labor is substantially smaller than his estimates.

[17] This difference between the translog and Cobb-Douglas models in terms of R-squared and precision of estimates does not arise in Moretti's across metropolitan area models.

For each industry-zip code cell, we also use the decennial Census data to calculate the share of the workforce that is white, black, Hispanic, or Asian-American. Using these shares, we calculate the average exposure of workers in an industry-zip code cell to workers of the same race at other firms in this PUMA (zpindexraceace). We calculate a similar measure for exposure of a firm's workforce to college educated workers of the same race in the PUMA in which their firm is located (zpindexracehedu). We then interact these two variables with the PUMA employment density and the PUMA share college educated, respectively, in order to test whether returns to agglomeration in terms of actual firm productivity depends upon the firm employees' within-race interaction opportunities. We also include direct controls for the racial composition of the workers in each firm cell.

The estimates for the Cobb-Douglas and translog models including these variables are shown in columns 3 and 4. For Cobb-Douglas, the estimates on the interactions between our spillover variables and our exposure indices are very noisy and insignificant. For example, while the coefficient on the share college interaction is the unexpected sign, the t-statistic is only 0.30. However, the use of the translog specification substantially increases the precision of our estimates of the spillover variables, the exposure variables and their interactions. The standard errors for the estimates of three variables related to employment density decrease by between 23 and 31 percent, while the decrease for the variables associated with share human capital is between 57 and 69 percent. The changes in R-Squared are very similar to the changes for the baseline model, and the Cobb-Douglas specification is soundly rejected. For the translog model, we find a strong statistically significant relationship between firm workers' average exposure to own race workers in PUMA and the return to employment density. In fact, our estimates suggest that there is no return to employment density for a firm whose workers have zero average exposure to own race workers. We do not find a statistically significant relationship between firm average exposure to own race college educated workers and returns to share college educated workers in a PUMA. However, the coefficient estimate is in the expected direction and sizable. The estimated return to share college with zero average exposure is less than 65 percent of the estimate in column 1.

In columns 5 and 6 we include controls for the unobserved ability of workers at the firm based on the residential location of those workers[18] and a control for whether we were able to match zip codes between the establishment and decennial Census data or were required to match based on industry by PUMA cells. In these last two models, the effect of the firm's own race match with its work location on return to density is very stable, and the effect of race match on return to share college increases by 40 and 20 percent. Further, the estimated returns to share college with zero average exposure are now 24 and 31 percent of the original estimate in column 1, respectively. Again, the failure to find significance on share college educated is not entirely surprising because manufacturing has lower returns than many other industries from exposure to share college educated and the racial differences and own race effect in return to share college is smaller in manufacturing than in the sample as a whole.

In terms of magnitudes, the effect of a one standard deviation increase in firm worker's average exposure to own race workers or to college educated workers evaluated at the average of workplace employment density and workplare share human capital is associated with a decline in firm total factor productivity of 0.6 percentage points for the own race density effect and 0.3 percentage points for the own race share college effect. We also can calculate racial differences

[18] The estimates for mean tract FE/unobserved worker ability are not shown because the variable is included in the translog production as another input and so involves several interactions. However, we also estimated Cobb-Douglas model with this control and find that as expected the mean tract FE variable has a positive and statistically significant effect on firm net revenue with an estimate of 0.145 and a t-statistic of 2.69.

in exposure to firm total factor productivity associated with these estimates. Specifically, by weighting the sample of firms by the number of workers of each group, we can calculate group average exposure to firm average own race or own race college educated exposure levels. Using racial differences in these numbers, we find that a one-standard deviation increase in employment density and share college implies a … decrease, respectively, in racial differences in exposure to firm total factor productivity.

VI. Discussion

This paper demonstrates that blacks receive lower returns to agglomeration economies in their place of work than do whites, a pattern that may contribute to overall racial income disparities and a host of other social concerns in the U.S. that are believed to be exacerbated by income inequality. Racial differences in both the return to employment density and human capital spillovers associated with worker education levels are robust to controlling for differences in the returns to agglomeration over demographics, industry, occupation and metropolitan area. The racial differences observed are substantially larger than the estimates on all other demographics except for education (which is smaller but in some cases comparable in magnitude), while education has a dramatically larger direct effect on wages than does race. These differences also cannot be explained away by controlling for the possibility that blacks make different residential, workplace and commuting mode decisions.

Several pieces of evidence suggest that black undercompensation is driven by race-specific social networks in the workplace. First, the returns to both agglomeration and human capital externalities increase as the share of workers who share an individual's race or ethnicity increases, and controlling for own race share of workers eliminates and/or significantly reduces the undercompensation of blacks. Second, when we split our sample by high/low return to agglomeration or high/low human capital spillover industries, we find that our results for share own race mirror our results for race. The racial differences in return and the effect of own race in workplace on return are both largest in the high return industries, which makes sense if our racial differences are driven by lower productivity gains from spillovers when blacks have fewer same race peers in work locations. Finally, we estimate a model of firm total factor productivity for a sample of manufacturing establishments to directly test whether the exposure of firm workers to workers of the same race at other firms affects firm productivity. We find strong evidence that the return to employment density rises as the average exposure of workers in a firm to same race peers rises. The pattern for the return from human capital spillovers is very similar and the effect is substantial in magnitude, but not statistically significant.

References

Ananat, Elizabeth. 2011. The Wrong Side(s) of the Tracks: The Causal Effects of Racial Segregation on Urban Poverty and Inequality. *American Economic Journal: Applied Economics*, 2011, 3(2): 34-66.

Audretsch, David B. and Feldman, Maryann P. 2004. Knowledge Spillovers and the Geography of Innovation. In J.V. Henderson and J.F. Thisse (Eds.) *Handbook of Urban and Regional Economics, Volume 4*. New York: North Holland.

Audretsch, David B. and Feldman, Maryann P. 1996. R&D Spillovers and the Geography of Innovation and Production. *American Economic Review*, 86, 630-640.

Bandiera, Oriana, Iwan Barankay and Imran Rasul (2005). Social Preferences and the Response to Incentives: Evidence from Personnel Data, *Quarterly Journal of Economics* 120(3): 917-962.

Bayer, Patrick, Stephen L. Ross, and Giorgio Topa. 2008. Place of Work and Place of Residence: Informal Hiring Networks and Labor Market Outcomes. *Journal of Political Economy*, 116, 1150-1196.

Bokenblom, Mattias and Kristin Ekblod. 2007. Sickness Absence and Peer Effects -Evidence from a Swedish Municipality. Working Paper.

Card, David and Jesse Rothstein. 2007. Racial Segregation and the Black-White Test Score Gap, *Journal of Public Economics*,

Ciccone Antonio and Robert E. Hall. 1996. Productivity and the Density of Economic Activity. *American Economic Review*, 86, 54-70.

Combes, Pierre-Philipi, Gilles Duranton, Laurent Gobillon. 2008. Spatial Wage Disparities: Sorting Matters! *Journal of Urban Economics*, 63, 723-742.

Cutler, David and Edward Glaeser. 1997. Are Ghettos Good or Bad? *Quarterly Journal of Economics*, 112, 827-872.

Di Addario, Sabrina and Eleanora Patacchini. 2008. Wages in the Cities: Evidence from Italy. *Labour Economics, 15, 1040-1061*.

Damm, Anna Piil. 2006. ‘‘Ethnic Enclaves and Immigrant Labour Market Outcomes: Quasi-Experimental Evidence.’’ Working Paper.

De Paola, Maria (2008). Absenteeism and Peer Interaction Effects: Evidence from an Italian Public Institute. Working Paper.

Duranton, Giles and Diego Puga. 2004. Micro-Foundations of Urban Agglomeration Economies. In J.V. Henderson and J.F. Thisse (Eds.) *Handbook of Urban and Regional Economics, Volume 4*. New York: North Holland.

Duranton, Giles and Diego Puga. 2001. Nursery Cities: Urban Diversity, Process Innovation and the Life-Cycle of Products. *American Economic Review*, 91, 1454-1477.

Dustman, Christian, Albrecht Glitz, and Uta Schonberg (2009). Job Search Networks and Ethnic Segregation in the Workplace. Working Paper.

Edin, Per-Anders, Peter Fredriksson, and Olof A°slund. 2003. Ethnic Enclaves and the Economic Success of Immigrants—Evidence from a Natural Experiment. Quarterly Journal of Economics 118(1):329–57.

Ellison, Glen, Edward Glaeser, and William Kerr. 2010. What Causes Industry Agglomeration? Evidence from Coagglomeration Patterns. *American Economic Review* 100, 1195-1213.

Falk, Armin and Andrea Ichino. 2006. Clean evidence on Peer Effects. *Journal of Labor Economics* 24: 39-57.

Feldman, Maryann P. and David B. Audretsch. 1999. Innovation in Cities: Science-based Diversity, Specialization, and Localized Competition. *European Economic Review*, 43, 409-429.

Fu, Shihe. 2007. Smart Café Cities: Testing Human Capital Externalities in the Boston Metropolitan Area. *Journal of Urban Economics*, 61, 86-111.

Fu, Shihe and Stephen L. Ross. 2010. Wage Premia in Employment Clusters: How Important is Worker Heterogeneity? Working Paper

Glaeser, Edward L. and David C Mare. 2001. Cities and Skills. *Journal of Labor Economics*, 19, 316-42.

Granovetter, Mark S. (1995), *Getting a Job: A Study of Contacts and Careers*, Cambridge, MA: Harvard University Press.

Henderson, J. Vernon. 2003. Marshall's Scale Economies. *Journal of Urban Economics*, 53, 1-28.

Hellerstein, Judith K., Melissa McInerney, and David Neumark. 2009. Neighbors And Co-Workers: The Importance Of Residential Labor Market Networks. NBER Working Paper #14201.

Hellerstein, Judith K., David Neumark, and Melissa McInerny. 2008. Spatial Mismatch or Racial Mismatch. *Journal of Urban Economics* 64: 464-469.

Hellerstein, Judith K., David Neumark, and Kenneth R. Troske. 1999. Wages, Productivity, and Worker Characteristics: Evidence from Plant Level Production Functions and Wages Equations. *Journal of Labor Economimcs* 17, 209-447.

Ihlanfeldt, Keith R. and David L Sjoquist. 1998. The Spatial Mismatch Hypothesis: A Review of Recent Studies and Their Implications for Welfare Reform. *Housing Policy Debate*, 9, 849-890.

Ioannides, Yannis M. and Linda Datcher Loury. 2004. Job Information Networks, Neighborhood Effects, and Inequality. *Journal of Economic Literature*, 42, 1056-93.

Kain, John F. 1992. The Spatial Mismatch Hypothesis: Three Decades Later. *Housing Policy Debate* 3(2):371–460.

Kain, John F. 1968. Housing Segregation, Negro Employment, and Metropolitan Decentralization. *Quarterly Journal of Economics*, 82, 175-197.

Lang, Kevin and Michael Manove. 2006. Education and Labor Market Discrimination. NBER Working Paper #12257.

Lindbeck, A., Palme, M. and Persson, M., (2007), "Social Interaction and Sickness Absence", *IFN Working Paper*, 725.

Mas, Alexandre and Enrico Moretti (2006). "Peers at Work," NBER Working Paper #12508.

Massey, Douglas and Nancy Denton. 1993. *American Apartheid: Segregation and the Making of the Underclass*, Cambridge, MA: Harvard University Press.

Moretti, Enrico. 2004. Human Capital Externalities in Cities. In J.V. Henderson and J.F. Thisse (Eds.) *Handbook of Urban and Regional Economics, Volume 4*. New York: North Holland.

Moretti, Enrico. 2004. Workers' Education, Spillovers, and Productivity: Evidence from Plant Level Production Functions. American Economic Review, 94, 656-690.

Nanda, Ramana and Jesper B. Sorenson. 2006. Peer Effects and Entrepreneurship. Working Paper.

Rosenthal, Stuart and William Strange. 2004. Evidence on the Nature and Sources of Agglomeration Economies. In J.V. Henderson and J.F. Thisse (Eds.) *Handbook of Urban and Regional Economics, Volume 4*. New York: North Holland.

Rosenthal, Stuart and William Strange. 2003. Geography, Industrial Organization, and Agglomeration. *Review of Economics and Statistics*, 85, 377-393.

Ross, Stephen L. Social interactions within cities: Neighborhood environments and peer relationships. 2011. In *Handbook of Urban Economics and Planning* (Eds. N. Brooks, K. Donaghy, G. Knapp). Oxford University Press.

Wheaton, William C. and Mark J. Lewis. 2002. Urban Wages and Labor Market Agglomeration. *Journal of Urban Economics*, 51, 542-62.

Wheeler, Christopher H. 2001. Search, Sorting and Urban Agglomeration. *Journal of Labor Economics*, 19, 879-99.

Wilson, William J. 1987. *The Truly Disadvantaged: The Inner City, the Underclass, and Public Policy*. Chicago: University of Chicago Press.

Yankow, Jeffrey J. 2006. Why Do Cities Pay More? An Empirical Examination of Some Competing Theories of the Urban Wage Premium. *Journal of Urban Economics*, 60, 139-61.

Table 1: Descriptives

	White	Black	Hispanic	Asian
Sample size	1,705,058	226,173	264,880	135,577
Dependent Variable				
Average hourly wage	28.6959 (45.6694)	19.5287 (31.1947)	17.7986 (32.7494)	26.1993 (40.8411)
Workplace Controls				
PUMA Employment density in 1000's/square KM	3.1611 (12.3109)	4.2424 (13.7415)	3.7645 (13.4660)	5.7770 (16.9347)
Share of college educated workers in PUMA	0.3773 (0.0982)	0.3775 (0.1013)	0.3566 (0.0985)	0.4090 (0.1052)
Employment density in own one digit industry	0.4606 (2.1408)	0.5348 (2.0841)	0.4436 (1.9072)	0.7810 (2.6242)
Share workers with college degree in industry	0.3549 (0.1703)	0.3456 (0.1687)	0.2959 (0.1580)	0.3926 (0.1726)
Share of workers of own race or ethnicity	0.7403 (0.1414)	0.1949 (0.1282)	0.2138 (0.1523)	0.1152 (0.0882)
Share college educated workers own race/ethnicity	0.3055 (0.0846)	0.0484 (0.0386)	0.0328 (0.0334)	0.0618 (0.0491)
Metropolitan Area Controls				
Percent college educated in MSA and occupation	0.0414 (0.0433)	0.0276 (0.0357)	0.0224 (0.0314)	0.0386 (0.0404)
Percent college educated in MSA and industry	0.0401 (0.0322)	0.0409 (0.0353)	0.0334 (0.0290)	0.0459 (0.0339)
Individual Worker Controls				
Age 30 to 39	0.4111 (0.4920)	0.4499 (0.4975)	0.5462 (0.4979)	0.4738 (0.4993)
Age 40 to 49	0.3663 (0.4818)	0.3605 (0.4801)	0.3103 (0.4626)	0.3455 (0.4755)
Age 50 to 59	0.2225 (0.4160)	0.1896 (0.3920)	0.1435 (0.3505)	0.1807 (0.3848)
Less than high school degree	0.0512 (0.2205)	0.1257 (0.3315)	0.3908 (0.4879)	0.1068 (0.3089)
High school degree	0.2043 (0.4032)	0.2863 (0.4520)	0.2181 (0.4130)	0.1159 (0.3201)
Associates degree	0.3020 (0.4519)	0.3560 (0.4788)	0.2391 (0.4265)	0.2037 (0.4027)
Four year college degree	0.2670 (0.4424)	0.1536 (0.3605)	0.0932 (0.2907)	0.2897 (0.4536)
Master degree	0.1126 (0.3161)	0.0546 (0.2272)	0.0324 (0.1770)	0.1706 (0.3762)
Degree beyond Masters	0.0629 (0.2428)	0.0239 (0.1528)	0.0264 (0.1603)	0.1132 (0.3168)
Single with no children	0.2296 (0.4206)	0.2811 (0.4496)	0.1822 (0.3860)	0.1483 (0.3554)
Married with no children	0.0289 (0.1674)	0.0762 (0.2653)	0.0744 (0.2624)	0.0276 (0.1638)
Single with children	0.3022 (0.4592)	0.2686 (0.4432)	0.2343 (0.4236)	0.2828 (0.4504)
Married with children	0.4393 (0.4963)	0.3741 (0.4839)	0.5091 (0.4999)	0.5413 (0.4983)
Born in the United States	0.9279 (0.2587)	0.8490 (0.3580)	0.3778 (0.4848)	0.1153 (0.3194)
Not born in U.S. resident less than 8 years	0.0149 (0.1212)	0.0272 (0.1626)	0.0966 (0.2954)	0.1807 (0.3848)
Not born in the U.S. resident 8 years or more	0.0572 (0.2322)	0.1238 (0.3294)	0.5256 (0.4993)	0.7040 (0.4565)
Additional Worker Variables				
mass transit dummy	0.0492 (0.2163)	0.1154 (0.3195)	0.0826 (0.2753)	0.0891 (0.2850)
work in central city	0.4257 (0.4946)	0.5602 (0.4964)	0.4845 (0.4998)	0.5077 (0.5000)
live in central city	0.2707 (0.4443)	0.5385 (0.4985)	0.4647 (0.4988)	0.3961 (0.4891)
share black in tract	0.0590 (0.1033)	0.4844 (0.3459)	0.1069 (0.1579)	0.0798 (0.1234)
work in high spillover industries	0.2842 (0.4510)	0.2310 (0.4215)	0.2180 (0.4129)	0.2736 (0.4458)

Note: Means and standard deviations are for a sample of 2,343,092 observations containing all male full-time workers aged 30 to 59 in the metropolitan areas with populations over 1 million residents where full-time work is defined as worked an average of at least 35 hours per week. Standard deviations are shown in parentheses.

Logarithm of the Wage Rate

Independent Variables	MSA FE
Employment density in own one digit industry	0.0118 (16.23)
Share workers with college degree in Industry	0.3894 (27.87)
African-American worker	-0.1465 (-45.11)
Hispanic worker	-0.1656 (-49.39)
Asian and Pacific Islander worker	-0.1349 (-24.00)
Other race	-0.1516 (-22.61)
Age 40-49	0.1010 (66.72)
Age 50-59	0.1568 (66.91)
Less than high school degree	-0.1456 (-59.85)
Associates degree	0.0851 (54.37)
Four year college degree	0.2711 (113.63)
Master degree	0.3903 (105.64)
Degree beyond Masters	0.5069 (117.4)
Single with children	0.0548 (22.19)
Married with children	0.2110 (94.37)
Married without children	0.1335 (96.46)
Not born in U.S. resident less than 8 years	-0.2533 (-46.21)
Not born in the U.S. resident 8 years or more	-0.0987 (-33.62)
Percent college educated in MSA and occupation	0.7453 (5.37)
Percent college educated in MSA and industry	1.1029 (8.23)
R-squared	0.2873

Note: OLS regressions. From column 1 of dindempdenbase.csv in disclosure folder.

Table 3: Baseline Agglomeration Model by Race or Ethnicity

White (sample size 1,705,058)	MSA FE
Employment density in 1000's per square KM	0.0138 (14.98)
Share workers with college degree	0.4390 (28.69)
R-squared	0.2461
African-American (sample size 226,173)	
Employment density in 1000's per square KM	0.0047 (5.80)
Share workers with college degree	0.1453 (6.57)
R-squared	0.2108
Hispanic (sample size 264,880)	
Employment density in 1000's per square KM	0.0097 (9.09)
Share workers with college degree	0.2069 (9.15)
R-squared	0.2536
Asian (sample size 135,577)	
Employment density in 1000's per square KM	0.0076 (6.48)
Share workers with college degree	0.3885 (12.19)
R-squared	0.3316

Note: OLS results, industry results are from column 1 of dindempdenbyrace.csv in disclosure folder.

Table 4: Model of the Wage Return to Agglomeration

Independent Variables	Employment Density	Share College Educated
African-American worker	-0.0060*** (-3.13)	-0.1680*** (-5.73)
Hispanic worker	-0.0014 (-0.57)	-0.0394 (-0.85)
Asian and Pacific Islander worker	-0.0013 (-0.50)	-0.0062 (-0.10)
Age 40-49	0.0011 (1.14)	0.0288* (1.92)
Age 50-59	0.0016 (1.09)	0.0566** (2.39)
Less than high school degree	-0.0048 (-1.51)	0.0401 (1.04)
Associates degree	-0.0003 (-0.17)	0.0501*** (2.72)
Four year college degree	-0.0016 (-1.07)	0.1466*** (6.55)
Master degree	0.0020 (1.19)	0.1419*** (4.52)
Degree beyond Masters	0.0054** (2.14)	0.1310** (2.14)
Single with children	-0.0018 (-0.28)	0.0367 (0.55)
Married with children	0.0034*** (2.92)	0.0440** (2.32)
Married without children	0.0022 (1.49)	0.0317 (1.40)
Not born in U.S. resident less than 8 years	-0.0037 (-1.02)	-0.0656 (-0.74)
Not born in the U.S. resident 8 years or more	-0.0023 (-1.17)	0.0080 (0.17)
R-square	0.0024	0.0028
Sample size	1465919	1465919

Note: basic slope models, from dindempslopemodelt4.csv

Table 5: Race Coefficients with varying Fixed Effects Structur

Variables	Metropolitan Area Fixed Effect	Tract Fix Effect	Block Group Fixed Effect	Tract-Cell Fixe Effect	Tract-Cell-Industry Fixed Effect	Tract-Cell-Blue Collar-Spillover
Race Coefficients from Wage Equation						
African-American worker	-0.1465	-0.0696	-0.0623	-0.0694	-0.0666	-0.0628
	(-45.11)	(-29.27)	(-26.20)	(-20.49)	(-6.75)	(-11.68)
Hispanic worker	-0.1657	-0.0939	-0.0859	-0.0660	-0.0621	-0.0563
	(-49.39)	(-39.59)	(-36.65)	(-17.45)	(-5.70)	(-8.95)
Asian and Pacific Islander worker	-0.1349	-0.1041	-0.1010	-0.0963	-0.0753	-0.0762
	(-24.00)	(-28.87)	(-27.91)	(-15.92)	(-3.98)	(-7.29)
Sample size	2343092	2343092	2343092	2343092	2343092	2343092
Race Differences in the Return to Agglomeration						
African-American worker	-0.0063	-0.0066	-0.0066	-0.0061	0.0020	-0.0028
	(-7.34)	(-7.64)	(-7.49)	(-3.17)	(0.40)	(-0.54)
Hispanic worker	-0.0012	-0.0017	-0.0022	-0.0014	0.0064	0.0063
	(-1.31)	(-1.78)	(-2.38)	(-0.56)	(1.11)	(0.98)
Asian and Pacific Islander worker	-0.0048	-0.0007	-0.0006	-0.0013	0.0022	0.0003
	(-4.22)	(-0.66)	(-0.59)	(-0.5)	(0.37)	(0.05)
Race Differences in the Return to Human Capital Externalities						
African-American worker	-0.1644	-0.1086	-0.1044	-0.1674	-0.1793	-0.1496
	(-10.53)	(-6.78)	(-6.47)	(-5.70)	(-2.71)	(-2.00)
Hispanic worker	-0.1268	-0.0502	-0.0428	-0.0423	-0.0668	-0.1266
	(-7.00)	(-2.73)	(-2.34)	(-0.92)	(-0.59)	(-0.99)
Asian and Pacific Islander worker	-0.1685	-0.0997	-0.0969	-0.0101	0.2423	0.2972
	(-6.94)	(-4.01)	(-3.91)	(-0.16)	(1.77)	(1.99)
sample size	2343092	2342887	2341593	1465267	527687	448936
Fraction of Sample in Singleton Cell	0	0.0001	0.0006	0.3708	0.7612	0.6037

Note: panel 1 from dindempdenracecells.csv; panels 2 and 3 from dindempslopevariousFE.csv.

Table 6: Analysis of Observations in Singleton Cell

Variables	Tract-Cell	Tract-Cell-Industry	Tract-Cell-Blue Collar-Spillover
First Stage Singleton Cell Coefficient			
Wage Equation	-0.0153	-0.0432	-0.026
	(-15.29)	(-39.62)	(-26.48)
Density Equation	-0.0144	-0.0707	-0.0383
	(-4.55)	(-14.35)	(-10.70)
Share College Equation	-0.00003	0.0142	-0.0011
	(-0.21)	(-53.92)	(-7.40)
Sample size	2343092	2342887	2342887
Second Stage Singleton Cell Interations			
Employment Density	-0.0008	-0.0027	-0.0011
	(-1.32)	(-3.71)	(-1.67)
Share College	0.0058	0.0578	-0.0077
	(0.52)	(6.28)	(-0.72)
Sample size	2342887	2342887	2342887

Note: from dindslopesingletondumy.csv; t statistics are in parentheses.

Table 7: Models of the Agglomeration and College Share Slopes by Worker Attributes

Variables	Employment Density	Share College Educated
Mass Transit Users		
African-American worker	-0.0083*** (-3.23)	-0.1096* (-1.92)
Hispanic worker	0.0011 (0.32)	0.0375 (0.25)
Asian and Pacific Islander worker	-0.0008 (-0.22)	-0.0535 (-0.29)
Sample size	75729	75729
Automobile Users		
African-American worker	-0.0025 (-1.02)	-0.1832*** (-6.11)
Hispanic worker	-0.0037 (-1.14)	-0.0692 (-1.45)
Asian and Pacific Islander worker	-0.0024 (-0.73)	-0.0220 (-0.34)
Sample size	1318372	1318372
Workplace in Central City		
African-American worker	-0.0065*** (-3.16)	-0.1902*** (-4.87)
Hispanic worker	-0.0002 (-0.09)	-0.0482 (-0.80)
Asian and Pacific Islander worker	0.0001 (0.05)	-0.1003 (-1.21)
Sample size	629272	629272
Workplace in Suburbs		
African-American worker	0.00002 (0.003)	-0.1667*** (-4.44)
Hispanic worker	-0.0039 (-0.86)	-0.0290 (-0.48)
Asian and Pacific Islander worker	-0.0033 (-0.73)	0.0862 (1.05)
Sample size	836647	836647
Tract Share Black above Metropolitan Area Average		
African-American worker	-0.0075** (-2.13)	-0.2175*** (-4.81)
Hispanic worker	-0.0050 (-0.97)	0.0543 (0.59)
Asian and Pacific Islander worker	-0.0044 (-0.60)	0.0679 (0.51)
Sample size	220492	220492
Tract Share Black below Metropolitan Area Average		
African-American worker	0.0022 (0.46)	-0.1122* (-1.65)
Hispanic worker	0.0002 (0.07)	-0.0871 (-1.62)
Asian and Pacific Islander worker	-0.0006 (-0.19)	-0.0392 (-0.54)
Sample size	1245427	1245427
Central City Resident		
African-American worker	-0.0056** (-2.24)	-0.1876*** (-3.84)
Hispanic worker	0.0028 (0.93)	-0.0543 (-0.72)
Asian and Pacific Islander worker	0.0008 (0.20)	-0.0729 (-0.72)
Sample size	410098	410098
Suburban Resident		
African-American worker	-0.0003 (-0.10)	-0.1503*** (-4.00)
Hispanic worker	-0.0063 (-1.36)	-0.0362 (-0.61)
Asian and Pacific Islander worker	-0.0032 (-0.81)	0.0417 (0.52)
Sample size	1055821	1055821

Table 7 continued

Workplace Employment Density above Metropolitan Area Average		
African-American worker	-0.0070*** (-3.33)	-0.1903*** (-4.86)
Hispanic worker	-0.0011 (-0.42)	-0.0166 (-0.26)
Asian and Pacific Islander worker	-0.0003 (-0.09)	-0.0332 (-0.4)
Sample size	612410	612410
Workplace Employment Density below Metropolitan Area Average		
African-American worker	-0.0029 (-1.11)	-0.1619*** (-4.46)
Hispanic worker	-0.0003 (-0.1)	-0.0711 (-1.28)
Asian and Pacific Islander worker	-0.0029 (-0.81)	0.0150 (0.18)
Sample size	853509	853509
Workplace Share College Educated above Metropolitan Area Average		
African-American worker	-0.0073*** (-3.52)	-0.1447*** (-4.22)
Hispanic worker	-0.0014 (-0.54)	-0.0653 (-1.21)
Asian and Pacific Islander worker	-0.0017 (-0.64)	-0.0426 (-0.60)
Sample size	863542	863542
Workplace Share College Educated below Metropolitan Area Average		
African-American worker	-0.0005 (-0.16)	-0.2097*** (-5.36)
Hispanic worker	-0.0005 (-0.15)	-0.0112 (-0.19)
Asian and Pacific Islander worker	0.0020 (0.04)	0.0390 (0.44)
Sample size	602377	602377
Worker in High Agglomeration Industry		
African-American worker	-0.0179*** (-2.92)	-0.2698*** (-3.46)
Hispanic worker	0.0026 (0.42)	-0.2198** (-2.19)
Asian and Pacific Islander worker	0.0010 (0.12)	0.0577 (0.37)
Sample size	230906	230906
Worker in Low Agglomeration Industry		
African-American worker	-0.0054*** (-2.80)	-0.1530*** (-5.09)
Hispanic worker	-0.0022 (-0.87)	-0.0059 (-0.12)
Asian and Pacific Islander worker	-0.0015 (-0.56)	-0.0190 (-0.29)
Sample size	1235013	1235013
Worker in High Spillover Industry		
African-American worker	-0.0066** (-2.40)	-0.3170*** (-5.63)
Hispanic worker	-0.0015 (-0.44)	-0.1545* (-1.66)
Asian and Pacific Islander worker	-0.0024 (-0.44)	-0.1137 (-0.82)
Sample size	264387	264387
Worker in Lower Spillover Industry		
African-American worker	-0.0052** (-2.31)	-0.1301*** (-4.11)
Hispanic worker	-0.0012 (-0.43)	-0.0191 (-0.39)
Asian and Pacific Islander worker	-0.0002 (-0.05)	0.0209 (0.31)
Sample size	1201532	1201532

Table 8: Agglomeration Model with Own Share Controls

Baseline Model	Employment Density	Share College Educated
African-American worker	-0.0061*** (-3.16)	-0.1678*** (-5.72)
Hispanic worker	-0.0015 (-0.60)	-0.0401 (-0.87)
Asian and Pacific Islander worker	-0.0013 (-0.50)	-0.0066 (-0.10)
with Own Share Controls		
African-American worker	-0.0002 (-0.08)	0.0624 (1.21)
Hispanic worker	0.0049* (1.66)	0.1917*** (3.01)
Asian and Pacific Islander worker	0.0060* (1.86)	0.1882** (2.57)
Own Share in Workplace	0.0151*** (3.93)	
Own Share College Educated		0.3756*** (5.44)
R-Square	0.0024	0.0029
Sample size	1465567	1465567

Note: baseline model results from dindheduownrace.csv columns B and F, should be the same to table 4, but slightly different because small difference in observations. Not sure why second panel from dindheduownrace.csv columns C and G.

Table 9: Own Share Controls by Returns to Agglomeration and Human Capital Spillovers

	Employment Density	Share College Educated
Worker in	*High Agglomeration Industry*	*High Spillover Industry*
African-American worker	-0.0069 (-0.87)	0.0092 (0.08)
Hispanic worker	0.0145* (1.75)	0.1697 (1.20)
Asian and Pacific Islander worker	0.0146 (1.46)	0.1858 (1.11)
Own Share in Workplace	0.0277** (2.20)	
Own Share College Educated		0.5617*** (3.27)
R-Square	0.0029	0.006
Sample size	230821	264332
Worker in	*Low Agglomeration Industry*	*Low Spillover Industry*
African-American worker	-0.0002 (-0.08)	0.0934* (1.67)
Hispanic worker	0.0036 (1.20)	0.2043*** (3.02)
Asian and Pacific Islander worker	0.0051 (1.54)	0.2075*** (2.69)
Own Share in Workplace	0.0136*** (3.43)	
Own Share College Educated		0.3594*** (4.88)
R-Square	0.0024	0.002
Sample size	1234731	1201220

Note： column 1 is from downraceowrkattsplits.csv columns R and S.
column 2 is from downraceheduworkattsplits.csv columns R and S.
high agglomeration industries, =1 for inducode 3,12,13

Table 10: Agglomeration Model with Own Share Controls - Manufacturing Subsample

Baseline Model	Employment Density	Share College Educated
African-American worker	-0.0129 (-0.40)	-0.1008 (-0.95)
Hispanic worker	-0.0068 (-0.33)	0.0662 (0.4)
Asian and Pacific Islander worker	0.0173 (0.79)	0.2368 (1.28)
R-Square	0.0022	0.0019
Sample size	169543	169543
with Own Share Controls		
African-American worker	-0.0034 (-0.09)	0.0692 (0.43)
Hispanic worker	0.0015 (0.06)	0.2317 (1.15)
Asian and Pacific Islander worker	0.0247 (0.97)	0.3391* (1.66)
Own Share in Workplace	0.0201 (0.63)	
Own Share College Educated		0.2510 (1.35)
R-Square	0.0022	0.0020
Sample size	169527	169527

panel 1 is from dmanufactslopemodel.csv

panel 2 is from dmanufactslopeownracemodel.pdf columns B and E.

Table 11 Total Factor Productivity models

Variables	Cobb-Douglas Model	Translog model	Cobb-Douglas Interaction Model	Translog Interaction Model	Interaction model with mean tract FE	Interaction model with dummy for missing zipcode data
Employment Density	0.0277*** (14.81)	0.0288***(21.29)	0.0114 (0.60)	-0.0028 (0.21)	-0.0011 (0.10)	-0.0019 (0.17)
Share College	0.2023*** (5.35)	0.2033*** (8.24)	0.2391 (1.43)	0.1311 (1.85)	0.0481 (0.67)	0.0640 (0.89)
Race Exposure Index			-0.1554 (1.28)	0.0032 (0.04)	0.0014 (0.02)	-0.0005 (0.01)
Density*Index			0.0437 (0.77)	0.0962*** (2.53)	0.0948*** (2.90)	0.0945*** (2.92)
Race/Coll Exp Index			0.0196 (0.13)	-0.0767 (0.71)	-0.0927 (0.89)	-0.0808 (0.77)
Density*Coll Index			-0.0815 (0.30)	0.1204 (1.14)	0.1699 (1.61)	0.1443 (1.37)
Zip Code Missing						-0.0248*** (5.52)
R Squared	0.8417	0.9085	0.8418	0.9086	0.9087	0.9088
Sample size	111695	111695	111695	111695	111538	111538

from tfpmodels2011disclose.csv columns G, I, K, L.