Human Capital and Economic Opportunity:
A Global Working Group

Working Paper Series

Working Paper No. 2011-027

Wage Premia in Employment Clusters: How Important is Worker Heterogeneity?

Shihe Fu

Stephen Ross

October, 2011

Human Capital and Economic Opportunity Working Group
Economic Research Center
University of Chicago
1126 E. 59th Street
Chicago IL 60637
humcap@uchicago.edu

Wage Premia in Employment Clusters: How Important is Worker Heterogeneity?

Shihe Fu
Southwestern University of Finance and Economics and
Cheung Kong Graduate School of Business, China

Stephen L. Ross
University of Connecticut

Key Words: Agglomeration, Wages, Sorting, Locational Equilibrium, Human Capital Externalities

JEL Codes: R13, R30, J24, J31

Any opinions and conclusions expressed herein are those of the author(s) and do not necessarily represent the views of the U.S. Census Bureau. All results have been reviewed to ensure that no confidential information is disclosed. Support for this research at the Boston and New York RDC from NSF (ITR-0427889) is also gratefully acknowledged.

Corresponding Author: Stephen L Ross, Department of Economics, University of Connecticut, 341 Mansfield Road U-63, Storrs, CT 06269-1063, stephen.l.ross@uconn.edu, 860-486-3533.

Wage Premia in Employment Clusters: How Important is Worker Heterogeneity?

Abstract

This paper tests whether the correlation between wages and the spatial concentration of employment can be explained by unobserved worker productivity differences. Residential location is used as a proxy for a worker's unobserved productivity, and average workplace commute time is used to test whether location-based productivity differences are compensated away by longer commutes. Analyses using confidential data from the 2000 Decennial Census Long Form find that the agglomeration estimates are robust to comparisons within residential location and that the estimates do not persist after controlling for commuting costs suggesting that the productivity differences across locations are not due to productivity differences across individuals.

Introduction

The strong correlation between wages and the concentration of economic activity has often been cited as evidence of agglomeration economies, but this correlation may also arise because highly productive workers prefer locations with high levels of economic activity. In this paper, a standard wage model is used to test for wage premia in agglomerated locations, except that a worker's residential location is used as a proxy for his or her unobservable productivity, under the premise that workers sort across residential locations based in part on their permanent incomes or innate labor market productivity. Further, in a locational equilibrium, identical workers should receive equal compensation, and therefore similar workers facing the same housing prices should receive the same wage net of commuting costs. The conceptual experiment is to compare two observationally equivalent individuals who reside in the same location and work in locations with different levels of agglomeration. Does the individual that works in the high agglomeration location earn a higher wage suggesting higher productivity at that work location, and if so does he or she also have a sufficiently longer commute so that the two workers receive the same real wage suggesting that the workers indeed have similar innate labor market productivity?

A central feature of most models of agglomeration economies is that agglomeration raises productivity. Since firms pay workers the value of their marginal production in competitive labor markets, a natural test for agglomeration economies is whether firms pay a wage premium in areas with concentrated economic activity.[1] Glaeser and Mare (2001), Wheeler (2001), Combes,

[1] Studies of agglomeration use a wide variety of approaches including examining productivity (Ciccone and Hall, 1996; Henderson, 2003), employment (Glaeser et al., 1992; Henderson, Kuncoro, and Turner, 1995), establishment births and relocations (Carlton, 1983; Duranton and Puga, 2001; Rosenthal and Strange, 2003), co-agglomeration of industries (Ellison, Glaeser, and Kerr, In Press; Dumais, Ellison, and Glaeser, 2002), product innovation (Audretsch and Feldman, 1996; Feldman and Audretsch, 1999) and land rents (Rauch, 1993; Dekle and Eaton, 1999). Also 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.

Duranton, and Gobillon (2004), Rosenthal and Strange (2006), Yankow (2006), Fu (2007) and Di Addario and Patacchini (2008) all find that wages are higher in large labor markets with high concentrations of employment. Many of these studies also find a positive link between wages and the human capital level associated with an employment concentration.[2]

A classic question in this literature is whether productivity is intrinsically higher in locations with a high concentration of employment, or whether high quality workers have simply sorted into those areas.[3] Glaeser and Mare (2001), Wheeler (2001), Yankow (2006) and Combes, Duranton, and Gobillon (2008) find evidence of an urban wage premium using longitudinal data, but worker fixed effects do explain a substantial portion of the raw correlation between employment concentration and wages. These studies often find that wages grow faster in larger urban areas, potentially due to faster accumulation of human capital.[4] The obvious limitation of this approach is that the relationship between agglomeration and wages is identified by the small fraction of people who move from one metropolitan area to another and those moves likely occur in response to attractive, potentially unexpected opportunities.[5]

Our paper proposes a new strategy that avoids relying on movers by drawing explicitly on several well-established features of urban economies. First, a worker's residential location is used as a proxy for his or her unobservable productivity attributes. Specifically, the paper

[2] Other studies, 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.

[3] Another major concern in the agglomeration literature is that individual places may have unobservables that contribute to higher productivity and so attract a concentration of economic activity so that high place specific productivity contributes to agglomeration rather than the other way around (Henderson, 2003; Ciccone and Hall, 1996). Regardless, most wage based studies of agglomeration focus on bias from sorting of workers across workplaces. In the context of this second concern, our analysis might be considered a test of worker sorting versus place specific productivity differences defined more broadly.

[4] The most compelling evidence behind the human capital accumulation story is provided by Glaeser and Mare (2001) who find that workers who migrate away from large metropolitan areas retain their earnings gains.

[5] In a cross-sectional study, Di Addario and Patacchini (2008) argue that they have identified causal effects of agglomeration on wages because there is almost no migration, i.e. sorting, across the labor markets covered in their sample of workers in Italy. The paper provides strong evidence that workers in large labor markets in Italy are more productive, but it is unclear whether this higher productivity arises from agglomeration economies or other unobservables, such as across market differences in the quality of the education system or attitudes towards work.

estimates wage premia across work locations that are located in the same metropolitan area[6] and examines whether these work location wage premia are robust to the inclusion of residential location fixed effects. This research design draws on the commonly accepted premise that individuals sort over residential locations based on tastes, which are partially unobservable and correlated with worker productivity.[7] For example, workers with higher productivity know that they can expect a higher lifetime income, and therefore these workers are likely to have a greater willingness to pay for neighborhood amenities. Workers residing in similar quality locations should have similar levels of productivity, and after controlling for residential location those workers should earn similar wages, unless their respective employment locations create productivity differences between the workers.[8]

Further, equilibrium in an urban economy requires that equivalent workers should obtain the same level of utility even if they live or work in different locations. After controlling for commuting time differences, workers residing in the same neighborhood should be indifferent between jobs in different locations, even if one of those locations contributes to higher productivity and therefore higher nominal wages. Rational workers will sort into locations with higher wages until congestion increases commuting time eroding the real value of the high nominal wage. In equilibrium, wage differences across locations must be entirely compensated

[6] Rosenthal and Strange (2006) also examine agglomeration effects on wages within metropolitan areas, but their primary focus is on the attenuation of these economies over space.

[7] A huge literature documents the fact that households are stratified across neighborhoods in part based on income. Gabriel and Rosenthal (1999) directly examine the effect of household sorting on wage models, Bayer, McMillan and Rueben (2004) estimate models of household sorting over neighborhoods based on race and income, and Epple and Sieg (1999) estimate models of household sorting over communities based on income.

[8] As will be discussed later, under specific assumptions, the residential fixed effects meet the conditions for a control function for our wage equation, see Blundell and Dias (2009). This strategy is also similar to an approach developed by Dale and Kruger (2002) in their study of higher education who condition on the set of schools to which students applied and were either accepted or rejected, and among students with similar choices and outcomes on this margin the selection into a specific school is assumed to be exogenous to quality of that school.

by longer commutes,[9] and unexplained location wage premia should not persist in models that control for both residential location and commute time unless those premia were created by unobserved productivity differences between workers. Specifically, a zero estimate on work location agglomeration in a model of wages net of commuting costs is consistent with no conditional correlation between agglomeration and worker unobserved productivity. While this compensation logic has been applied in the quality of life literature (Roback, 1982; Gyouko, Kahn, and Tracy, 1999; Albouy, 2008, 2009) and in Davis, Fisher, and Whited (2009) and Glaeser and Gottlieb (2009) to study wage premia across metropolitan areas, this logic has not been exploited to examine agglomeration economies within metropolitan areas, even though within a metropolitan area job and residential mobility rates are substantially higher than across metropolitan mobility (Ross, 1998).

We draw a sample of individuals residing in mid-sized to large metropolitan areas from the confidential data of the long form of the 2000 U.S. Decennial Census and estimate the relationship between the concentration of employment in their workplace (employment location) and their wage, controlling for a standard set of individual controls plus occupation, industry, and metropolitan area fixed effects. We find agglomeration effects that are comparable in size to earlier estimates, as well as evidence that the wages are higher in locations with more educated workers.[10] The agglomeration estimates are unchanged by the use of residential location fixed effects to control for unobserved worker productivity differences, and our estimates suggest that a one standard deviation increase in agglomeration as measured by total employment raises log

[9] Timothy and Wheaton (2001) examine the capitalization of commutes into wages within urban labor markets Some earlier studies of urban wage gradients include Madden (1985), Ihlanfeldt (1992), McMillen and Singell (1992) and Ihlanfeldt and Young (1994).

[10] The influence of the presence of educated workers on wages is discussed in the context of human capital externalities. However, this paper does not make any explicit attempt to test the various competing hypotheses concerning the underlying causes of agglomeration economies. See Ellison, Glaeser, and Kerr (In Press) and Fu (2007) for recent work on this question.

wages by 0.033, which is approximately half of the across metropolitan wage premium. This share of the cross-sectional urban wage premium is comparable to the 60 to 80 percent reduction found by Glaeser and Mayer (200) and the 50 percent reduction found by Combes, Duranton, and Gobillon (2008) when they control for individual fixed effects. The robustness of our agglomeration estimates to the inclusion of residential fixed effects is consistent with the small estimated within metropolitan area correlation between agglomeration and our observable measure of productivity, education.[11] Further, commute time can explain most of the relationship between the agglomeration variable and wages with very reasonable values on total commuting costs of less than 1.8 times the wage, suggesting that the estimated agglomeration effect is not due to unobserved worker productivity. Similar findings arise for human capital externalities using an extended model that controls for the average education level in the work location.

The two obvious weaknesses of this approach are that residential location may provide an imperfect control for unobserved worker quality and that workers may sort over commute time based on their unobservables creating a correlation between commutes and worker productivity.[12] Concerning imperfect neighborhood controls, we extend our basic model to allow for sorting on factors other than permanent income. By directly calculating the bias using an errors-in-variables framework (see appendix), we demonstrate that the inclusion of residential fixed effects reduces bias in our agglomeration estimates, leads to attenuation of the estimated coefficients on observed human capital, and the magnitude of the bias reduction is quite sensitive to the attenuation of human capital estimates. Empirically, we examine the estimated coefficients

[11] The across metropolitan area correlation between education and agglomeration is substantially larger than the within metropolitan correlation, suggesting a substnatial across metropolitan correlation between ability and agglomeration, which is consistent with the large declines in the agglomeration estimates found by Glaeser and Mayer (200) and Combes, Duranton, and Gobillon (2008) from the inclusion of individual fixed effects.

[12] The systematic selection of workers across commutes based on income or wage rate is well established in urban economics, see LeRoy and Sonstelie (1983) and Glaeser, Kahn, and Rappaport (2008).

on the education variables and find that the estimates are attenuated by the inclusion of the residential controls, exactly as is expected if the residential controls are capturing worker productivity unobservables. Attenuation increases substantially as residential controls are refined to smaller geographic units to capture more unobservables, and yet our agglomeration estimates are very stable suggesting little bias from worker heterogeneity in the original OLS estimates.[13] In addition, our results are robust in models that drop all individual covariates, which should exacerbate bias if imperfect sorting is a serious concern.

Concerning the commute time model, we directly test whether workers sort across commutes based on observable measures of human capital. We find that the conditional correlation between average workplace commute time and worker education is between 0.019 and 0.034, and these small correlations are associated with no appreciable attenuation of the human capital coefficients from the inclusion of commute time as a control. After controlling for other model variables, workers are not sorting across commutes based on observable measures of human capital, which is supportive of the maintained assumption that workers are not sorting over commutes based on unobservable ability.[14] Further, using the errors-in-variable calculations, we demonstrate that the small agglomeration estimates in the net of commute wage model provide an upper bound for the bias in the fixed effect estimates, as long as the estimate

[13] One might reasonably ask whether this attenuation could be explained by measurement error in our education variables, given the common perception that measurement error is exacerbated by the inclusion of fixed effects. The answer is yes and no. The attenuation bias from measurement error is only exacerbated by the inclusion of fixed effects when the fixed effects can systematically explain variation in the control variable, in this case our observable measure of productivity - education. Therefore, one must ask why the residential fixed effects are correlated with observable productivity, presumably sorting, and then ask whether the fixed effects should not be also correlated with unobservable aspects of productivity. Therefore, while some of the attenuation in parameter estimates may be due to increased attenuation from measurement error in education, this attenuation likely can only arise due to a correlation between residential location and productivity variables and so supports our claim that the increased attenuation is evidence that our fixed effects provide a proxy for productivity in wage regressions.

[14] Altonji, Elder, and Tabor (2005) suggest that the degree of selection on observables may provide a good indication of the potential selection on and bias from unobservables. Further, given the anticipated strong correlation between education and ability, sorting over commutes based on ability would likely show up as a correlation between commutes and education.

on commute time is at or below the true value. Our model estimates provide substantial evidence of agglomeration economies for quite conservative values of commuting costs.

In summary, we apply our identification strategy to a large, representative sample and estimate the relationship between concentrated employment and wages using a broad population of workers residing in mid-sized and large U.S. metropolitan areas. Even after conditioning on residential location, we find estimates of agglomeration economies that are comparable in magnitude to traditional estimates. Further, the empirical relationship between agglomeration and net of commute wages is far too small to explain our agglomeration estimates suggesting again that they are not seriously biased by unobserved worker productivity. Therefore, we conclude that location specific wage premia associated with agglomeration within metropolitan areas cannot be explained by worker heterogeneity.

The paper is organized as follows. The next section presents our empirical methodology and summarizes our errors-in-variables analysis.[15] The third and fourth sections describe the data and the findings, and the fifth section concludes.

Methodology

The basic empirical model is quite similar to models investigated in previous wage studies of agglomeration economies where it is assumed that firms pay workers their marginal revenue product and so differences in nominal wages capture the returns to higher productivity arising in agglomerated locations. The logarithm of individual *i*'s wage (y_{ij}) in location *j* is

$$y_{ij} = \beta X_i + \gamma Z_j + \alpha_i + \varepsilon_{ij}, \qquad (1)$$

where X_i is a vector of individual observable attributes, Z_j is employment concentration in the employment location *j*, α_i is an individual specific random effect that captures heterogeneity in

[15] See appendix for the complete errors-in-variables analysis.

labor market productivity, but is uncorrelated with X_i, and ε_{ij} is a random error that allows an individual's current earnings or wage to differ from their permanent income or earnings capacity, possibly due to the idiosyncratic match between workers and jobs.[16] If individuals sort over employment locations based on their expected wage or inherent productivity $(\beta X_i+\alpha_i)$, or tastes that are correlated with productivity, the unobserved component of productivity α_i will be correlated with Z_j or

$$E[Z_j \alpha_i] \neq 0,$$

biasing estimates of γ. Typically, the concern is that high ability individuals sort into high agglomeration locations biasing the estimates of agglomeration effects on wages upwards. For example, see Gabriel and Rosenthal (1999), Bayer, McMillan and Rueben (2004) and Epple and Sieg (1999) for evidence of individuals and households systematically sorting across neighborhoods and communities based on wages or income.

Residential Location as a Proxy for Worker Unobservables

Our proposed solution to this problem is based on the simple idea that individuals sort into residential locations based on their unobservables, and therefore one can minimize unobservable differences between workers by comparing individuals who reside in the same location. Specifically, under assumptions specified below, the residential location fixed effects will act as a control function for worker productivity. Specifically, Blundell and Dias (2009) formally define a control function δ as $(\varepsilon_{ij}, \alpha_i) \perp (Z_j, X_i) \mid \delta$ based on notation in equation (1)

[16] The assumption that X_i and α_i are uncorrelated can be made without loss of generality by considering β as representing the reduced form relationship between observables and wages. Specifically, let κ_i be the true unobserved productivity that correlates with X_i and assume that the conditional expectation of κ_i can be written as a linear function λX_i. Under those conditions, the expectation of equation (1) may be written as follows $E[y_{ij} - \gamma Z_j \mid X_i] = \beta X_i + \kappa_i = \beta X_i + E[\kappa_i \mid X_i] + (\kappa_i - E[\kappa_i \mid X_i]) = (\beta + \lambda) X_i + \alpha_i$ yielding a reduced form model specification where $E[\kappa_i \mid X_i]$ represents that bias λ and α_i is orthogonal to X_i by construction because κ_i in the last term has been differenced by its conditional expectation.

so that conditional on δ OLS will yield consistent estimates of γ. The properties of residential sorting models with taste unobservables have been well established by Epple and Platt (1998), Epple and Sieg (1999), and Bayer and Ross (2006). Specifically, these models imply perfect stratification so that if individuals sort across residential locations based solely on a common measure of location quality (W_k) and their demand for location quality, then each residential location k will contain workers in a continuous interval of location quality demand.

If we assume demand depends on permanent income based on a worker's innate productivity ($\beta X_i+\alpha_i$), worker productivity will be monotonic in location quality, or in other words locations can be ordered so that if

$$W_k < W_{k+1}$$

for location k then in equilibrium

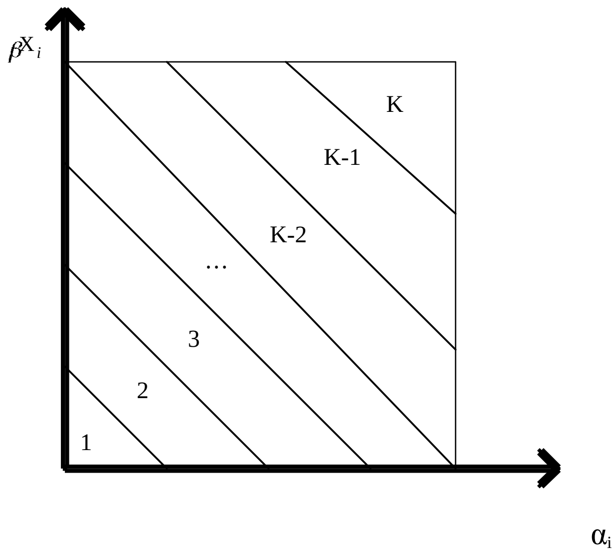
$$\delta_k < \beta X_i + \alpha_i < \delta_{k+1}$$

for all individuals i residing in location k where δ_k is assumed to be less than δ_{k+1} for any k. If there are a large number of residential choices then

$$\delta_k \approx \beta X_i + \alpha_i \qquad (2)$$

Figure 1 illustrates this partial equilibrium sorting pattern where a band of individuals with similar permanent income $\beta X_i + \alpha_i$ reside in the same community and these groups are monotonically ordered by permanent income over K communities of increasing attractiveness. The slanted lines represent loci of boundary individuals who all have the same permanent income and are indifferent between the neighborhoods on either side of a locus.

Figure 1. Sorting Equilibrium



Under these assumptions, δ_k satisfies the definition of a control function, and consistent estimates of γ can be obtained by substituting equation (2) into equation (1) and estimating the following equation

$$y_{ijk} = \delta_k + \gamma Z_j + \varepsilon_{ijk}\,, \tag{3}$$

where δ_k might be captured by a vector of residential location fixed effects. In this specification, workers in the same residential location are assumed to have identical productivity, and so unexplained wage differences across workers in the same residential location must reflect aspects productivity associated with the work location, such as agglomeration economies or some innate aspect of productivity associated with agglomerated locations, rather than worker unobservables.

A Test for the Correlation between Worker Unobservables and Agglomeration

The second component of our strategy for testing whether the estimated value of γ is biased by unobserved differences in worker productivity draws upon the locational equilibrium requirement that no workers desire to change either their residential or employment locations. As discussed earlier, observationally equivalent workers residing in the same location should earn the same wages net of commute or the same real wage unless some workers have higher productivity based on unobservables. Under the assumption that the urban economy is in a locational equilibrium, we attribute any systematic differences in wages net of commuting costs to the sorting of individuals across work locations based on individual productivity unobservables. A finding of no systematic relationship between real wages and agglomeration in a model that controls for commuting costs is consistent with a zero correlation between unobserved differences in worker productivity and agglomeration, and therefore consistent with unbiased estimates of agglomeration economies in the model of nominal wages.

Formally, locational equilibrium requires that

$$U(y_j, P_k, V_{jk}) = U(y_{j'}, P_k, V_{j'k}) \qquad (4)$$

where U is the indirect utility function of a type of individuals who reside in location k and are observed in both employment locations j and j', P_k is the price per unit of housing services in location k, and V_{jk} is the commuting time or cost between locations k and j. Fujita and Ogawa (1982) and Ogawa and Fujita (1980) consider a simple model of the urban economy with production externalities (agglomeration economies) and commuting where work hours and land consumption are fixed. In this model, the equilibrium condition in equation (4) requires that wages net of commuting costs must be the same across all employment locations j conditional on a worker's residential location. Specifically,

$$U(y_j - \eta V_{jk}, P_k) = U(y_{j'} - \eta V_{j'k}, P_k) \text{ or } y_j - \eta V_{jk} = y_{j'} - \eta V_{j'k} \qquad (5)$$

over all work locations j and j' where η is the per mile or minute commuting costs.[17] The reader should note that wages net of commute costs or real wages in this context are constant across

[17] See Ross (1996) and Ross and Yinger (1995) for examples of the same locational equilibrium condition in a traditional monocentric urban model with an exogenous city center. In those papers, housing demand is endogenous, and the locational equilibrium condition in equation (5) still arises. In fact, this equation will hold and commute time is monetized in any model where either leisure does not enter preferences or total work hours including commute time are fixed.

work locations even though agglomeration economies exist as reflected by nominal wage differences across work locations.

Building on the logic of this model, we will specify wage equations in which wages compensate workers for commute costs in a work location, as opposed to wages being based on worker's marginal product in a location.[18] Given the sorting described in equation (2), workers in the same residential location have the same innate productivity or permanent income and so should receive the same real wages. Wages for individuals residing in residential location *k* and working in location *j* can be written as

$$y_{ijk} = \delta_k + \eta t_{jk} + \xi_{ijk}\,, \tag{6}$$

where t_{jk} is the commute time, η captures the monetary value of all commuting costs including time spent communing, and ξ_{ijk} is a stochastic error term. This model captures compensation of workers as opposed to the productivity of workers as modeled in equation (3), and ξ_{ijk} represents unobservable factors that affect the utility associated with individual *i* in work location *j*, again potentially arising from the idiosyncratic match between a worker and job, but in this case in terms of how attractive the worker finds the job. A comparison of equations (3) and (6) implies that

$$\gamma Z_j = \eta t_{jk} + (\xi_{ijk} - \varepsilon_{ijk})\,. \tag{7}$$

Equation (7) suggests that the influence of agglomeration on wages should be completely captured by commuting time. If agglomeration has no influence on wages after controlling for commuting costs, workers in the same residential location are receiving equivalent

[18] Gabriel and Rosenthal (1996) and Petitte and Ross (1999) apply similar logic to empirically study the welfare impacts of residential segregation by testing whether African-Americans had longer commutes after including residential location fixed effects, and in the case of Petitte and Ross (1999) also including employment location fixed effects, as controls for housing price and wage differentials that might compensate for longer commutes.

compensation, which could only occur in a locational equilibrium if those workers have equivalent productivity. On the other hand, if agglomeration explains wages net of commuting costs, those wage differentials (presuming they persist in equilibrium) must represent individual workers being compensated for their innate ability, which would suggest that agglomeration estimates continue to be biased by worker sorting on unobservables even after controlling for residential location fixed effects.

Admittedly, commute time and agglomeration will be highly correlated in equilibrium, and so workers by sorting into high agglomeration locations should also have sorted into work locations that require long commutes. So, the inclusion of commute time in the wage model may erode the coefficient on agglomeration even if the agglomeration coefficient is driven by workers sorting on unobserved ability. It is important to stress, however, that the test is not whether the addition of commute time as a control eliminates the agglomeration coefficient, but rather whether an effect of agglomeration on wages exists after conditioning on commuting costs at a reasonable valuation. While we will estimate models with both commute time and agglomeration controls in order to observe an estimated value of commuting costs,[19] the key test is whether the agglomeration coefficient is near zero when the commute time coefficient takes on a reasonable value for representing commuting costs, which can be assessed by setting the commute time coefficient to specific values based on outside information.

Imperfect Neighborhood Sorting

The assumption of complete sorting based on permanent income or innate ability implies that the residential location fixed effects fully capture individual productivity. Such a strong assumption seems unrealistic since residential location choice is influenced by tastes that

[19] An analysis of bias from errors in variables, located in the appendix, demonstrates that this model will provide an unbiased estimate of commuting costs under the assumption that our agglomeration proxy contains measurement error.

are unlikely to be perfectly correlated with permanent income, and in practice observed human capital variables, like education, have strong predictive power in our wage equations even after controlling for residential location fixed effects. The predictive power of human capital variables rejects the implications of equation (3).

Therefore, the empirical model is extended to consider the situation where the residential location fixed effect δ_k differs from the productivity of an individual residing in k by a random error (μ_i) that is uncorrelated with $\beta X_i+\alpha_i$ or

$$\delta_k = \beta X_i + \alpha_i + \mu_i \,. \tag{10}$$

For example, μ_i may represent individual tastes for neighborhood quality that are independent of productivity or permanent income. This heterogeneity leads to a classic errors-in-variables problem. This result is easily observed by substituting equation (10) into equation (1) yielding

$$y_{ijk} = \delta_k + \gamma Z_j + (\varepsilon_{ij} - \mu_i)\,, \tag{11}$$

where δ_k is positively correlated with μ_i by construction. The reader should note that μ_i represents tastes and only enters the wage equation because the fixed effect contains μ_i.

The negative correlation between the fixed effect δ_k and the error (ε_{ij} -μ_i) will attenuate the estimates of δ_k towards zero. Given the assumption that Z_j is positively correlated with worker ability ($\beta X_i+\alpha_i$), the estimate of γ continues to be biased upwards since worker ability is imbedded in the fixed effect and the associated correlation between Z_j and δ_k biases the coefficient on Z_j upwards. Intuitively, the attenuated fixed effect estimates provide only a partial control for $\beta X_i+\alpha_i$, and potentially the estimates might be improved by directly including X_i in the location fixed effect model specification

$$y_{ijk} = \beta X_i + \delta_k + \gamma Z_j + (\varepsilon_{ij} - \mu_i)\,. \tag{12}$$

Further, given that α_i is unobserved, the estimate of β, the coefficient vector for observable human capital, conditional on residential fixed effects will be attenuated relative to the OLS estimates from equation (1). As illustrated in Figure 1, two individuals with different X_i's residing in the same neighborhood or community are likely to have different α's; otherwise, they would have had different permanent incomes and chosen different neighborhoods. This selection process into neighborhoods creates a negative correlation between X_i and α_i within any residential location (Gabriel and Rosenthal, 1999; Bayer and Ross, 2006) attenuating the estimated coefficients on the human capital variables. In our case, however, this bias is an advantage because the predicted attenuation bias in the human capital coefficient estimates provides a metric for assessing whether the residential location fixed effects successfully capture variation associated with individual unobserved productivity. Specifically, the estimated coefficients on human capital variables in the residential fixed effects model can be compared to the estimates from a simple regression model without fixed effects, and if the inclusion of fixed effects reduces the estimated coefficients then the residential fixed effects have captured some variation associated with unobserved productivity attributes.[20]

The problem described above involves bias arising from errors-in-variables with multiple correlated regressors. Given the complexity of this problem, we turn to numeric calculations of the bias in estimated parameters in order to confirm the intuition discussed in the preceding paragraphs. Specifically, we manually calculate the formulas for the omitted bias in each parameter, use our data to estimate the variances and covariances for key observables, and then calculate the bias in our key parameters. The details of this analysis are shown in the appendix.

[20] Measurement error in the education variables will also cause attenuation bias, which might be exacerbated by the inclusion of residential fixed effects, but the measurement error bias is only worsened by the fixed effects if the residential fixed effects can explain productivity attributes. See the discussion in Footnote 13.

These calculations confirm the key assertions earlier in the section. The inclusion of residential fixed effects into a model that controls for observable productivity or human capital leads to a substantial reduction in the bias in the agglomeration estimates, and the inclusion of commute time dramatically reduces the estimates on the agglomeration variable. Further, if households do not sort across commutes based on their permanent income, the coefficient on the agglomeration variable after controlling for both residential location fixed effects and commutes is larger than the bias on the agglomeration estimates after controlling for residential location fixed effects and so provides an upper bound on the bias from imperfect sorting. We also confirm that the coefficient estimate on human capital attenuates with the inclusion of residential fixed effects due to unobserved worker productivity. Further, sensitivity analyses confirm that the reduction in bias is quite stable over parameter values except when the attenuation of the human capital estimate changes, which has large impacts on the reduction in bias.

The simulations also confirm our concerns that any correlation between productivity and commute time will bias our analysis. Specifically, the estimated coefficient on agglomeration in a model that controls for commute time may no longer provide an upper bound for the bias in the residential fixed effect agglomeration model. However, these calculations also indicate that the agglomeration coefficient after conditioning on commute time does provide an upper bound as long as the estimate on commute time is equal or below the parameter's true value. This result confirms our earlier intuition that zero estimates on agglomeration after controlling for actual commute costs suggests no bias from unobserved worker productivity variables in the residential fixed effect agglomeration estimates.

Sample and 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.[21] These restrictions lead to a sample of 2,234,092 workers.

The dependent variable, logarithm of wage rate, is based on a wage that is calculated by dividing an individual's 1999 labor market earnings by the product of number of weeks worked in 1999 and usual number of hours worked per week in 1999. The wage rate model includes a standard set of labor market controls including variables capturing age, race/ethnicity, educational attainment, marital status, presence of children in household, immigration status, as well as industry, occupation,[22] and metropolitan area fixed effects. Finally, the model includes controls for share of college-educated employees in a worker's industry or occupation at the metropolitan level.[23] The mean and standard errors for these variables are shown in Table 1 separately for the college educated and non-college educated subsamples.

We consider two alternative specifications to capture employment concentration: the number of workers employed in a employment location, which we will refer to in this paper as

---

[21] 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**.**

[22] Workers are classified into 20 major occupation codes and 15 major industry codes.

[23] These controls are similar in spirit to a control used by Glaeser and Mare (2001) for occupation education levels nationally. Obviously, the industry, occupation, and metropolitan area fixed effects even when combined with the metropolitan area industry and occupation education controls do not absorb as much variation as the MSA-occupation cell fixed effects used by Rosenthal and Strange (2006). Given our focus on models that control for the large number of residential tract fixed effects, it is not feasible to simultaneously include this large array of MSA-occupation fixed effects. However, the models without residential fixed effects have been re-estimated with MSA-occupation fixed effects and results were similar. Further, models including MSA-occupation fixed effects were estimated for some subsamples based on a small number of very large MSA's, where residential fixed effects could be included directly in the model rather than differenced. Again, all findings are robust.

the workplace, and the workplace employment density for a variety of workplace definitions.[24] Similarly, models are estimated controlling for residential location at a variety of levels of aggregation. Our preferred specification defines residential locations at the census tract level and workplaces at the residential Public Use Microdata Area (PUMA) level, where residential PUMAs are defined based on having a minimum of 100,000 residents, and measures agglomeration using workplace employment density. The control for commute time is based on the average commute time[25] for all full time workers employed at a workplace.[26] Additional specifications are estimated that control for the fraction of workers in the workplace PUMA who have a college degree or above. All standard errors are clustered by workplace.

Results

Table 2 presents the results for a baseline model of agglomeration economies in wages using both controls for total employment and employment density at the residential PUMA level. The estimates on the control variables are quite standard and stable across the two specifications. Based on these estimates, adding 10,000 workers to a workplace is associated with a 0.54 percent

[24] The agglomeration variables are constructed using all full time workers not just the prime-age, male workers present in the regression sample.

[25] In principle, the model should include a control for the commute of the marginal worker, but such information is not typically available. Timothy and Wheaton (2001) and Small (1992) describe the circumstances under which average commute time will be a sufficient statistic for marginal commute time, and Small (1992) provides empirical and simulation evidence suggesting that average commutes are a good proxy for marginal commutes.

[26] Since the models are identified based on within residential location variation, the workplace commute time implicitly controls for commute time between place of residence and place of work without the measurement error inherent in estimating average commute time between every residence to workplace combination. In principle, the appropriate way to handle such measurement error is to instrument for residence to workplace commute time with average workplace commute time, rather than simply including workplace commutes directly in the wage model. The IV estimates controlling for residence to workplace commute time are very similar in magnitude (slightly smaller) to the estimates presented here and discussed in this paper, and obviously the estimated coefficients on the agglomeration variables are unaffected by such a specification change.

increase in wages while an increase in employment density of 1000 workers per square kilometer is associated with a 0.24 percent increase in wages.[27]

Fixed Effect Estimates

Panel 1 of Table 3 contains the estimates for the baseline model, as well as the models that include residential location fixed effects at the census tract level and that include both residential fixed effects and average commute time at the workplace. In the residential location fixed effect model, the positive relationship between agglomeration and wages is robust to the inclusion of these controls, which should increase the similarity of individuals over which the effect of agglomeration economies is identified. In fact, including residential fixed effects has little impact on the estimated coefficients on agglomeration. The failure to find substantial bias from workers sorting on unobservables across work locations within metropolitan areas is consistent with the evidence of sorting on observable human capital variables. The within metropolitan area correlation between worker education level and employment density after controlling for other observables is quite small: 0.034 for our education index,[28] 0.029 for whether a worker has at least a four year college degree, and 0.019 for whether a worker has at least a high school degree or above.

Of course, one explanation for not finding evidence of sorting bias is that our residential location fixed effects do not successfully capture worker unobserved productivity variables. However, as discussed earlier, if the residential location fixed effects provide effective controls for individual productivity unobservables due to residential sorting, the coefficient estimates on

---

[27] Rosenthal and Strange (2006) estimate models using the Public Use Microdata Sample and controlling for total employment within spatial rings of employment estimated from workplace PUMA's. Our estimated magnitudes using total employment in actual workplace PUMA's are comparable to theirs.

[28] This index was created for the correlation estimates used for our errors-in-variables bias calculations presented in Tables A1 and A2 in the appendix, and the index is a linear combination of the educational attainment dummies based on the coefficient estimates on education presented in Table 3.

human capital should be biased towards zero by the inclusion of residential location fixed effects. We find such evidence of attenuation bias for both models. In the density model, the inclusion of residential fixed effects reduces the estimates on above master's degree, master's degree, four year college degree, associate degree, and high school diploma from 0.665, 0.546, 0.424, 0.225, and 0.138 to 0.511, 0.424, 0.330, 0.175, and 0.108, respectively, a reduction of between 22 and 23 percent in all coefficients.[29]

The magnitude of the within metropolitan area estimates of agglomeration economies are quite reasonable. The within metropolitan estimates are comparable in magnitude to simple OLS estimates arising from comparisons across metropolitan areas.[30] Specifically, we find that in a model controlling for standard individual attributes, a one standard deviation increase in metropolitan wide employment or employment density increases logarithm of wages by 0.062 and 0.044, respectively. Meanwhile, using the census tract fixed effects estimates, a one standard deviation in workplace total employment or density leads to an increase in the logarithm of wages of 0.033 and 0.034, which is between half and three-quarters of the traditionally estimated across metropolitan wage premium.

In addition, in panel 2 of Table 3, we examine a wage model that controls for the logarithm of the agglomeration variables converting the estimated effects to elasticities. The pattern of estimates in panel 2 is nearly identical to the pattern for the baseline estimates shown in panel 1 of Table 3, and the estimates imply that a doubling of agglomeration economies based on total employment or density is associated with a 4.3 and 2.0 percent increase in wages,

[29] Attenuation of estimates in the total employment model is virtually identical to attenuation in the employment density model.
[30] We estimate the same wage model controlling for metropolitan total employment or the metropolitan wide employment density, as well as regional fixed effects.

respectively, which bracket Combes, Duranton, and Gobillon's (2001) elasticity estimate of 3 percent after controlling for individual fixed effects in a sample of movers.[31]

Further, in panel 3, we examine the effect of increasing the bias from unobserved ability by restricting the number of individual controls. Specifically, we re-estimate the models in panel 1 dropping all individual covariates including the education, age and family structure variables, which correlate very strongly with labor market outcomes. Naturally, the R-squares of the estimated models fall substantially from 0.29 to 0.20 in the OLS model with the omission of these measures of human capital. However, the within metropolitan area OLS estimates of agglomeration economies are essentially unchanged at 0.054 and 0.0022 for total employment and employment density. The residential location fixed effects estimates increase somewhat from 0.051 to 0.058 for total employment and from 0.0026 to 0.0029 for employment density, which are relatively small increases given the omission of so much information relevant to labor market outcomes. These very stable estimates of agglomeration, when so much observable information has been excluded, is consistent with our finding that within metropolitan area agglomeration estimates are not substantially biased by workers sorting based on their unobservables.[32]

In addition, in panels 4 and 5, we examine the effect of basing our estimates on more homogenous comparisons. First, the sample is restricted to single, male workers. This population of workers is less likely to have their residential location decision influenced by marital and family obligations. The pattern of estimates is very similar. For example, both the OLS and residential fixed effects employment density estimates are 0.0018.[33] In panel 5, we organize the

[31] All other estimates in the paper involve employment and density levels rather than logs in order to be comparable to other recent work that uses the Census microdata to study wages and agglomeration economies.
[32] We also experimented with models that do not contain the industry and occupation fixed effects and the pattern and magnitude of estimates was again very similar.
[33] It is worth noting that the decline in estimated agglomeration effects for the sample of single, male workers is not driven by marital status. Rather, single male workers are younger and have less education on average than married males, and our estimated agglomeration effect increases moderately with an individual's level of human capital. In

sample into cells of observationally equivalent individuals based on discrete variables for age, race/ ethnicity, education, family structure, and immigration status,[34] and control for cell by census tract fixed effects so that our estimates are truly based on comparing very similar individuals who reside in the same location. As in panel 4, agglomeration estimates are not affected by the inclusion of residential location controls.

Commute Time or Compensation Models

Columns three and six of Table 3 contain the estimates for the model containing residential location fixed effects and workplace average commute time. The inclusion of commute time as a control eliminates most of the relationship between the agglomeration variables and wages, and the magnitude of the estimated coefficients fall by more than a factor of five in our baseline model, the estimated employment density coefficient falls from 0.0026 to 0.0004 (panel 1), and in the logarithm model our estimated effect falls from 2.0 percent to 0.5 percent (panel 2). The estimates fall by a factor of 10 in the model with no covariates (panel 3), in the sample of single men the estimate changes sign and becomes insignificant (panel 4), and when controlling for cell by tract fixed effects the estimates again fall by almost a factor of 10. The estimated agglomeration effects are almost completely compensated by longer commutes.

Next, the key question to ask is whether the commute time control is truly capturing compensation of identical productivity individuals for wage differentials across workplaces or whether commute time is acting as a proxy for unobserved productivity due to workers systematically sorting across commutes and/or workplaces. First, we can examine the extent of worker sorting based on observable measures of productivity. After conditioning on residential

addition, we estimated models for single and married workers separately by education level finding similar results that agglomeration economies increase with education levels for both single and married males.

[34] Households are divided by three age, five race, six education, four family structure based on presence of children by marital status, and three immigration categories based on whether born in the U.S. and time in the U.S. if not allowing for a total of 1,080 possible cells.

location fixed effects and other model variables, the correlation between commute time and worker education level is 0.019 for whether the worker has a high school degree, 0.029 for whether the worker has a college degree or years of education, and 0.034 for the education based wage index.[35] As suggested by Altonji, Elder, and Tabor (2005) in the context of Catholic schools, the conditional correlation between a variable of interest and observable measures of ability likely provides some indication of the conditional correlation between that variable and unobserved ability, and in our data we find a fairly small conditional correlation between average workplace commutes and education, our observable measure of human capital.

Most importantly, we examine whether the estimates on commute time are consistent with anticipated time, monetary, and any other disutility costs of commuting. If commute time were proxying for unobserved productivity variables rather than compensating for wage differentials, the coefficient estimate on commute time would likely exceed reasonable commuting costs in order to capture high unobserved worker productivity in agglomeration locations. However, if the estimate on commute time is reasonable, the agglomeration estimates in the commute time model captures the portion of agglomeration estimates that is not explained or compensated away by commutes and so might represent payments to workers based on their innate ability or other worker unobservables. In order to assess the magnitude of the commute time estimates, we start with a simple back of the envelope calculation using the estimates from panel 1 of Table 3. Specifically, a one minute increase in one way commute time leads to approximately 0.7 percent increase in wages on average. With an eight hour day, a two minute increase in round trip commutes represents 0.42 percent increase in the length of the workday. Dividing these numbers implies that a 0.7 percent point estimate is consistent with total

[35] Workplace commute time and the education variables are regressed on the PUMA fixed effects model in Table 4 except that the education dummy variables and the agglomeration variables are excluded from the model.

commuting cost including the value of time spent and monetary costs being compensated at 1.64 times the market wage.

For more precise estimates, we shift to an instrumental variables framework in which we control for an individual's time spent commuting as a share of average daily work time including commuting time (two way commute time divided by the sum of commute time and one-fifth of average hours worked per week assuming a five day work week) and use the average commute time for the workplace PUMA as an instrument.[36] This specification uses the same source of variation to identify the compensation of commutes, but uses the share of work time spent commuting in order to scale the effect and estimate compensation as a fraction of the wage rate. For example, if commuting increases the work day by one percent, the wages for time spent at work would need to increase by one percent in order to just compensate the worker at their wage for the time spent commuting.

The estimates for the total employment and employment density models in the first column of Table 4 are 1.78 and 1.82 suggesting that time spent commuting is compensated at less than double the wage rate, which is consistent with Timothy and Wheaton (2001) who find compensation rates of between 1.6 and 3.0 times the wage rate.[37] Further, Small (1992) estimates that on average the monetary cost of commuting is both proportional to and similar in magnitude to an individual's wage suggesting a compensation rate of two if people also value their time spent commuting at the wage rate and suggesting an even larger compensation rate if we recognize that monetary commuting costs are paid with after tax income. Finally, the next two columns present estimates that restrict the coefficient on commute time share to 1.5 and 1.0,

[36] The first stage includes all control variables in the log wage equation except for the agglomeration variable so that the entire effect of agglomeration is captured directly by the estimated coefficient on the agglomeration variable. Note that models in which the agglomeration variable is included in the first stage yield nearly identical results.
[37] Another factor to consider in evaluating these commute time costs is that commuting costs are typically paid using after tax income and our wage measures based on the census data are pre-tax.

respectively. The estimates on the agglomeration variables rise and are a little more than half the size of the Table 3 estimates when the commute time share coefficient is restricted to 1.0. These conservative estimates suggest that at least half of the estimated agglomeration economies cannot be compensated away and so cannot be due to unobserved productivity differences across individuals.[38]

Alternative Workplace and Residential Location Definitions

Table 5 presents estimates using two additional workplace definitions to measure employment density and commute time. As discussed above, the residential PUMA is defined to contain approximately 100,000 residents. The largest alternative definition is the workplace Public Use Microdata Areas (workplace PUMAs), which are often substantially larger than residential PUMAs especially near central cities and publically available,[39] but also quite idiosyncratic across metropolitan areas with some areas having almost as many workplace as residential PUMAs and others areas with millions of residents having only one or two workplace PUMAs. There are approximately 25 percent more residential PUMAs than workplace PUMAs in our sample. We also examine models where agglomeration and workplace commute are

[38] We also examine models where we control for or instrument with the average commute time between workplace and residential PUMA's. The estimates on commute time fall consistent with greater measurement error in the place to place commute time, and the agglomeration coefficients rise somewhat. The resulting IV commute time estimates are approximately 1.5 similar to the commute time parameter values that were used in Columns 2 and 5 of Table 4 and our agglomeration estimates are similar to the estimates in those columns, as well. Further, when we set the coefficient to 1.8, which we believe is a more reasonable value of commuting costs, our model using place to place commute time yields the same agglomeration estimate as observed in Table 4 using the workplace specific commute time. Similar results arise if we simply control for commuting time as a share of workday. Commute time estimates fall consistent with substantial measurement error, but if estimate models of wage net of commute cost using actual commute time and a coefficient of 1.8 we get agglomeration estimates very similar to

[39] Readers can find examples of similar estimations using data available in the Public Use Microdata sample of the U.S. Census in earlier working paper versions of this manuscript. See University of Connecticut Working Paper number 2007-26.

measured at the zip code, and our sample contains about six times as many zip codes as residential PUMAs. Residential fixed effects are included at the census tract level.[40]

The standardized estimates are the largest for residential PUMA suggesting that measurement error might be worse for larger or smaller workplace definitions, but the other estimates are still sizable and statistically significant. The pattern of results across columns is remarkably similar except for two minor differences. When workplace PUMA is used to measure agglomeration, the inclusion of fixed effects leads to an increase in the agglomeration estimate of approximately 17 percent. For both the residential PUMA and zip code models, agglomeration estimates are fairly stable changing by only about 8 percent with the inclusion of fixed effects. Further, the attenuation of the agglomeration estimate with the inclusion of commute time is much larger, typically a factor of 10, for the workplace PUMA definition. While these results might lead one to prefer workplace PUMA, we chose the more uniform residential PUMA as the baseline workplace definition in order to be conservative.

Table 6 presents estimates based on alternative geographic definitions of residential location. The largest neighborhood definition is the residential PUMA with estimates shown in panel 1, followed by estimates based on the smaller zip codes in panel 2. Census tracts are even smaller with populations between 1,500 and 8,000 (panel 3), and block groups are smaller still with populations between 600 and 3,000 (panel 4). The fixed effect estimates of agglomeration, as well as the fixed effect plus commute time model, are nearly identical across the four panels. However, the attenuation in the estimates on education variables, which indicates the degree to which the fixed effects can capture unobserved ability, varies dramatically. The inclusion of residential PUMA fixed effects leads to an attenuation of 8-12 percent, while zip codes lead to

[40] From this point forward, we only present estimates using employment density, but estimates using total employment are very similar.

15-17 percent, census tracts to 22-23 percent, and block groups to 24-26 percent declines in estimated education coefficients. The fixed effects capturing more detailed spatial resolution lead to greater attenuation presumably capturing a more homogeneous population on productivity in these smaller neighborhoods, and yet produce very similar agglomeration estimates, which based on our simulations is consistent with our finding that unobserved individual productivity variables do not bias within metropolitan estimates of agglomeration.[41]

Improving the Residential Location Controls

In this section, we consider expanded fixed effect models that might better control for unobserved heterogeneity. Ortalo-Mange and Rady (2006) find substantial heterogeneity among homeowners within neighborhoods, but considerable homogeneity among renters and among homeowners who moved into the neighborhood at similar times. Presumably, renters and recent homeowners chose this neighborhood based on current prices and neighborhood amenities and therefore are similar, while homeowners that moved to the neighborhood in earlier years chose this neighborhood based on different prices and amenity levels. Alternatively, one physical residential location might be divided into different submarkets based on the type of housing stock. For example, an individual who resides in a small loft in an apartment building may be very different from someone who selects a large single family dwelling in the same residential location, even if the two individuals have similar levels of observable human capital.

In order to address these concerns, we develop residential location fixed effects by tenure in residence and by housing stock categories. For the tenure of residence fixed effect model, a full set of tract fixed effects are created for each of the following categories: renters, owners who

[41] In principle, the greater attenuation in the education coefficients arising from the use of block groups fixed effects suggest that we should use block groups fixed effects for the models that follow, but we return to the use of census tract fixed effects for the rest of the paper in order to facilitate comparison to the earlier results. Regardless, the substantive results of the paper are robust to any of the geographies considered in this section.

have been residing in the neighborhood for less than one year, owners who have been residing in the neighborhood between one and five years, and owners who have been residing in the neighborhood for more than five years. For the housing stock model, tract fixed effects are created for each of seven housing stock categories: mobile home, multifamily 1 bedroom or less, multifamily 2 bedroom, multifamily 3 bedroom or more, single family 2 or less bedrooms, single family 3 bedrooms, and single family 4 ore more bedrooms. The results are shown in Table 7, and the expansion of the residential fixed effects has little impact on the estimated agglomeration effects. Further, as with more geographically narrow residential locations, both sets of controls significantly increase the attenuation of the coefficient estimates on the human capital variables, from between 22-23 to 26-29 percent, while not affecting the agglomeration estimates.[42]

In addition, the locational equilibrium test for agglomeration economies requires the assumption that individuals in the same residential location face the same price per unit of housing services. This assumption may not be reasonable because it is expensive and often prohibited by zoning to change the type of housing on specific parcels of land. As a result, the price per unit of housing services may vary considerably across different forms of housing in the same neighborhood due to differences between current demand and the historical supply of housing in this neighborhood. Our submarket fixed effects help address this concern, and the resulting commute time estimates and the impact of including commute time on agglomeration estimates are nearly identical to the results in Table 3.

Alternative Subsamples and Robust Commute Time Estimates

Table 8 presents estimates for a series of regional subsamples. The first panel presents results for the full sample with the subsequent panels containing the estimates for metropolitan

[42] In principle, one might wonder whether these different geographic definitions have different implications for different size and density metropolitan areas. However, our agglomeration estimates are very stable as we restrict our analysis to a smaller number of larger metropolitan areas.

areas in the Northeast, Midwest, South, and West regions. The qualitative findings concerning the coefficient estimate on employment density in Table 3 are replicated across all four regions. The estimated impact of agglomeration is quite stable when controlling for residential fixed effects and then falls to near zero after the inclusion of a control for commute time. The raw coefficient estimates on employment density exhibit substantial variation across regions, but in part this is due to different urban environments in each region. After standardizing the coefficients using the within metropolitan area standard deviation of employment density, the estimated agglomeration effects in the fixed effect models are closer in magnitude with values of 0.034, 0.055, 0.027, 0.015, and 0.017 for the full sample, Northeast, Midwest, South, and West regions, respectively. Also, the commute time results are quite stable across the samples with estimates ranging from 0.0066 to 0.0069 over the four regions as compared to 0.0069 for the full sample, again consistent with commuting costs that are noticeably less than twice the wage.[43]

Table 9 presents a similar set of estimates for subsamples based on college education, transportation mode, and race/ethnicity. Standardized estimates of agglomeration are substantially higher for the college educated, non-Hispanic white, and mass transit subsamples ranging between 0.037 and 0.053, as compared to a range of 0.022 to 0.025 for the non-college educated, minority, and automobile user subsamples. As in Table 8, the agglomeration estimates fall dramatically when commute time is included in the models, and the estimates on commute time are stable across the college educated, non-college educated workers, automobile commuters, and mass transit commuters with estimates ranging between 0.0064 and 0.0074. The only exception to this finding is the white-minority split, where the estimated relationship

[43] The standardized estimates on total employment for each region are quite similar to the density estimates presented in the paper.

between commute time and wages of 0.0034 for minorities is substantially smaller than the 0.0081 estimate for non-Hispanic white subsample.[44]

This last finding should not be surprising considering previous research concerning minority commutes and the spatial mismatch hypothesis. For example, Gabriel and Rosenthal (1996) and Petitte and Ross (1999) both find racial differences in commutes that cannot be compensated for by differences in housing prices and/or wages. Our findings are consistent with the notion that minorities are in a locational equilibrium when compared to each other, but are under compensated for their commutes when compared to the majority population. Barriers faced by minorities or other imperfections in the labor market may differentially affect minorities preventing them from being fully compensated for their commutes. For example, Hellerstein, Neumark, and McInerney (2008) find that access to African-American held jobs, rather than overall employment access, explains the employment outcomes of African-Americans.

Finally, our finding of larger agglomeration economies for college graduates is notable because it is consistent with Moretti (2009) who finds that high skill individuals have been migrating to more agglomerated, higher cost metropolitan areas. His evidence suggests that the reason behind this is a shift in the demand for labor in these areas and is not simply a stronger preference for large city amenities among the college educated. Similarly, we find that the agglomeration wage premium is higher for college educated individuals. The lower agglomeration coefficient for minorities may reflect the lower levels of educational attainment among minorities, while the large estimated agglomeration effects for the mass transit sample is likely due to the high concentration of mass transit users in the Northeast.[45]

[44] Again, the pattern of results is nearly identical in models using total employment to measure agglomeration.
[45] Northeast residents comprise more than half of the mass-transit subsample. The authors recognize that transportation mode choice is clearly endogenous to labor market earnings, and these models are estimated primarily to examine the stability of commute time coefficients across subsamples.

Workplace Human Capital Specification

Table 10 presents estimates for models that also include a control for the workplace share of workers with a four year college education or higher. The extended model is still consistent with agglomeration economies with a coefficient estimate of 0.0022 for the fixed effects model with the full sample (panel 1), very similar to the estimate in Table 3, much smaller estimates after controlling for commute time, and an estimated coefficient on commute time of 0.0066 consistent with reasonable commuting costs. The education level of workers in a workplace is also positively associated with wages, which is consistent with the standard human capital externalities explanation that often arises in this context (Rauch, 1993; Moretti, 2004; Rosenthal and Strange, 2006). As before, the estimated effect of agglomeration on wages is robust to the inclusion of residential fixed effects, but the estimated effects of share college educated decline from 0.359 to 0.151 when residential fixed effects are included.[46] These findings are consistent with the notion that high skill individuals sort into places with concentrations of highly educated workers. As with agglomeration, the coefficient on share college educated declines substantially (a factor of 3 in panel 1) with the inclusion of commute time as a control. Panels 2, 3 and 4 of Table 10 present estimates for a model with no covariates for the full sample, for the baseline model using the subsample of single, male workers, and for a model controlling for observationally equivalent individual cells by census tract fixed effects. As in Table 3, all results are robust, and the general pattern of findings persists.

[46] Rosenthal and Strange (2006) control separately for the number of college educated and non-college educated workers. They find that the number of college educated workers increases wages while the number of non-college educated workers reduces wages. While this result is fairly robust, the number of college and non-college workers in a workplace PUMA have correlations above 0.97 even after conditioning on metropolitan area or residential PUMA. Further, we have identified at least one specification where we observe a sign reversal so that wages fall with the number of college educated. When we estimate models that are directly comparable to Rosenthal and Strange (2006), our estimated effect sizes are fairly similar in magnitude to their estimates for a five mile radius circle.

Summary and Conclusions

We find that within metropolitan wage premia cannot be explained by high productivity workers sorting into agglomerated locations and so these wage premia must arise from location specific differences, either agglomeration or other location productivity differences. Specifically, the estimates for both total employment and employment density indicate a positive relationship between workplace agglomeration and firm wages, and these estimates are unchanged by the inclusion of residential location controls intended to absorb worker heterogeneity, even when residential fixed effects are included for each group of observationally equivalent individuals. The magnitudes of these estimates are sizable with standardized effects between one-half and three-quarters of the estimated across-metropolitan wage premium for the same sample. Estimates for the individual education variables attenuate when the residential controls are included, which is consistent with the residential controls capturing unobserved heterogeneity. The attenuation increases substantially as location controls are refined by focusing on smaller geographic measures of residential location or housing submarkets within residential locations, and these changes have no impact on the agglomeration estimates consistent with our main finding of no bias from worker sorting. This finding is also consistent with the small within metropolitan area correlation between agglomeration and observable human capital.

The inclusion of commute time dramatically reduces the estimated effect of agglomeration on wages. The estimates on commute time imply commuting costs of less than two times the wage, which is consistent with the current literature on commuting costs, and the correlation between observed measures of human capital and commute time is quite small. These findings suggest that the observed nominal wage differences do not represent differences in ability across workers because the commute time variable captures commuting costs accurately and wages net of commuting costs do not vary systematically across employment locations, presumably leaving similar workers with similar levels of well-being.

Further, bias calculations across a variety of parameter values indicate that the agglomeration estimates after controlling for commute time form an upper bound for the bias in the fixed effect agglomeration estimates, as long as estimates on commute time are not biased upwards. Even in the extreme case where we assume that the total commuting costs are only one time the wage, the implied causal effects of agglomeration are substantial, between one-quarter and three-eighths of the across-metropolitan wage premium. All findings including the implied commuting costs are robust across a wide variety of subsamples, different geographic definitions of workplace and residential neighborhood, use of housing submarket by neighborhood fixed effects, as well as a very challenging test for our estimation strategy where we omit all individual level covariates substantially increasing the variance attributable to unobserved worker ability.

Finally, an extended specification is estimated that includes a variable intended to capture human capital externalities, share of workers with a four year college degree or above. As in the previous literature, we find that wages increase with the concentration of college-educated workers. The effect of human capital externalities on wages falls by over half with the inclusion of fixed effects, likely because high productivity individuals are sorting across work locations based on education levels. However, the resulting fixed effect estimates are still sizable, and the inclusion of commute time substantially reduces the estimated relationship between wages and share college educated workers variable supporting our view that a substantial fraction of our fixed effect estimates represent the causal effect of human capital externalities on wages.

The results in this paper also have more general implications concerning the nature of urban economies. Only limited empirical evidence on urban wage gradients exists to support the idea that urban labor markets are in a locational equilibrium. This paper provides substantially more direct evidence by demonstrating that wage gradients can substantially compensate for nominal wage differences within metropolitan areas. Further, if agglomeration economies eventually plateau and possibly decline on the margin at very high concentrations of employment, empirical estimates of agglomeration effects may understate the total importance of agglomeration in urban economies, especially in cities with relatively effective transportation systems, because in equilibrium workers should continue to crowd into the high employment concentration locations until marginal productivity declines sufficiently to assure equal wages net of commuting costs.

References

Albouy, David. 2009. What are Cities Worth? Land Rents, Local Productivity, and the Value of Amenities. NBER Working Paper No.14981.

Albouy, David. 2008. Are Big Cities Really Bad Places to Live? Improving Quality-of-Life Estimates across Cities. NBER Working Paper No.14472.

Altonji, Joseph G., Todd E. Elder, Christopher R. Tabor. 2005. Selection on Observed and Unobserved Variables: Assessing the Effectiveness of Catholic Schools. *Journal of Political Economy*, 113, 151-184.

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.

Bayer, Patrick, Robert McMillan and Kim Rueben. 2004. An Equilibrium Model of Sorting in an Urban Housing Market. NBER Working Paper No. 10865.

Bayer, Patrick and Stephen L. Ross. 2006. Identifying Individual and Group Effects in the Presence of Sorting: A Neighborhood Effects Application. NBER Working Paper No. 12211.

Blundell, Richard and Monica C. Dias. 2009. Alternative Approaches to Evaluation in Empirical Microeconomics. *Journal of Human Resources*, 44, 565-640.

Brownstone, David and Kenneth A. Small. 2005. Valuing Time and Reliability: Assessing the Evidence from Road Pricing Demonstrations. *Transportation Research Part A*, 39, 279-293.

Carlton, Dennis W. 1983. The Location of Employment Choices of New Firms: An Econometric Model with Discrete and Continuous Endogenous Variables. *Review of Economics and Statistics*, 65, 440-449.

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.

Dale, Stacy B. and Alan B. Krueger. 2002. Estimating the Payoff to Attending a More Selective College: An Application of Selection on Observables and Unobservables. *Quarterly Journal of Economics*, 117, 1491-1527.

Davis, Morris, Jonas D.M. Fisher, and Toni M. Whited. 2009. Agglomeration and Productivity: New Estimates and Macroeconomic Implications. Working Paper.

Dekle, Robert and Jonathan Eaton. 1999. Agglomeration and Land Rents: Evidence from the Prefectures. *Journal of Urban Economics*, 46, 200-214.

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

Dumais, Guy, Glen Ellison, and Edward Glaeser. 2002. Geographic Concentration as a Dynamic Process. *Review of Economics and Statistics*, 84, 193-204.

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.

Ellison, Glen, Edward Glaeser, and William Kerr. In Press. What Causes Industry Agglomeration? Evidence from Coagglomeration Patterns. *American Economic Review*.

Epple, Dennis and Glen J. Platt. 1998. Equilibrium and Local Redistribution in an Urban Economy when Households Differ in Both Preferences and Incomes. *Journal of Urban Economics*, 43, 23-51.

Epple, Dennis and Holger Sieg. 1999. Estimating Equilibrium Models of Local Jurisdictions. *Journal of Political Economy*, 107, 645-681.

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.

Fujita, Masahisa and Hideaki Ogawa. 1982. Multiple Equilibria and Structural Transition of Nonmonocentric Urban Configurations. *Regional Science and Urban Economics*, **12**, 161–196.

Gabriel, Stuart and Stuart Rosenthal. 1999. Location and the Effect of Demographic Traits on Earnings. *Regional Science and Urban Economics*, 29, 445-461.

Gabriel, Stuart and Stuart Rosenthal. 1996. Commutes, Neighborhood Effects, and Earnings: An Analysis of Racial Discrimination and Compensating Differentials. *Journal of Urban Economics*, 40, 61-83

Glaeser, Edward L and Joshua Gottlieb. 2009. The Wealth of Cities: Agglomeration Economies and Spatial Equilibrium in the United States. NBER Working Paper #14806.

Glaeser, Edward L., Matthew Kahn, and Jordan Rappaport. 2008. Why Do the Poor Live in Cities. *Journal of Urban Economics*, 63, 1-24.

Glaeser, Edward L., Hedi D. Kallal, Jose A. Scheinkman, and Andrei Shleifer. 1992. Growth in Cities. *Journal of Political Economy*, 100, 1126-1152.

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

Gyourko, Joseph, Matthew Kahn and Joseph Tracy. 1999. Quality of Life and Environmental Comparisons. in E. Mills and P. Cheshire (Eds.) *Handbook of Regional and Urban Economics,* Vol 3. Amsterdam: North Holland.

Hellerstein, Judith K., David Neumark, Mellisa McInerney. 2008. Spatial Mismatch or Racial Mismatch? *Journal of Urban Economics*, 64, 464-479.

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

Henderson, J. Vernon, Ari Kuncoro, and Matt Turner. 1995. Industrial Development in Cities. *Journal of Political Economy*, 103, 1067-1085.

Ihlanfeldt, Keith R. 1992. Intraurban Wage Gradients: Evidence by Race, Gender, Occupational Class, and Sector. *Journal of Urban Economics,* 32, 70-91.

Ihlanfeldt, Keith R. and Madelyn V. Young. 1994. Intrametropolitan Variation in Wage Rates: The Case of Atlanta Fast-Food Restaurant Workers. *Review of Economics and Statistics*, 76, 425-433.

LeRoy, Stephen F. and Jon Sonstelie. 1983. Paradise Lost and Regained: Transportation Innovation, Income, and Residential Location. *Journal of Urban Economics,* 13, 67-89.

Madden, Janice F. 1985. Urban Wage Gradients: Empirical Evidence. *Journal of Urban Economics,* 18, 291-301.

McMillen, Daniel P. and Larry D. Singell. 1992. Work Location, Residence Location, and the Intraurban Wage Gradient. *Journal of Urban Economics*, 32, 195-213.

Moretti, Enrico. 2008. Real Wage Inequality. NBER Working Paper No.14370.

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.

Ogawa, Hideaki and Masahisa. Fujita. 1980. Equilibrium Land Use Patterns in a Non-monocentric City. *Journal of Regional Science*, 20, 455-475.

Ortalo-Mange, Francois and Sven Rady. 2006. Housing Market Dynamics: On the Contribution of Income Shocks and Credit Constraints. *Review of Economic Studies*, 73, 459-485.

Rauch, James E. 1993. Productivity Gains from Geographic Concentration off Human Capital: Evidence from the Cities. *Journal of Urban Economics,* 34, 380-400.

Petitte, Ryan A. and Stephen L. Ross. 1999. Commutes, Neighborhood Effects, and Compensating Differentials: Revisited. *Journal of Urban Economics,* 46, 1-24.

Roback, Jennifer. 1982. Wages, Rents, and the Quality of Life. *Journal of Political Economy*, 90, 1257-1278.

Rosenthal, Stuart and William Strange. 2006. The Attenuation of Human Capital Spillovers: A Manhattan Skyline Approach. *Journal of Urban Economics*, 64, 373-389.

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. 1998. Racial Differences in Residential and Job Mobility: Evidence Concerning the Spatial Mismatch Hypothesis. *Journal of Urban Economics*, 43, 112-136.

Ross, Stephen L. 1996. The Long-Run Effect of Economic Development Policy on Resident Welfare in a Perfectly Competitive Urban Economy. *Journal of Urban Economics*, 40, 354-380.

Ross, Stephen L. and John Yinger. 1995. Comparative Static Analysis of Open Urban Models with a Full Labor Market and Suburban Employment. *Regional Science and Urban Economics*, 25, 575-605.

Small, Kenneth A. 1992. *Urban Transportation Economics, Fundamentals of Pure and Applied Economics Series, Vol. 51*. New York: Harwood Academic Publishers.

Small, Kenneth A., Clifford Winston & Jia Yan. 2005. Uncovering the Distribution of Motorists' Preferences for Travel Time and Reliability. *Econometrica*, 73, 1367-1382.

Timothy, Darren and William C. Wheaton. 2001. Intra-Urban Wage Variation, Employment Location, and Commuting Times. *Journal of Urban Economics*, 50, 338-366.

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.

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.

Appendix

A Model with Imperfect Sorting and Measurement Error

The complete sorting and full compensation assumptions developed above have two implications that are inconsistent with the empirical data that will be used in this study. First, as discussed in the body of the paper, the model with complete sorting based on permanent income requires that the residential location fixed effects fully capture individual productivity. The empirical model is extended to consider the situation where the residential location fixed effect δ_k differs from the productivity of an individual residing in k by a random error (μ_i) that is uncorrelated with $\beta X_i+\alpha_i$ or

$$\delta_k = \beta X_i + \alpha_i + \mu_i . \tag{A1}$$

This heterogeneity leads to a classic errors-in-variables problem. This result is easily observed by substituting equation (A1) into equation (1) yielding

$$y_{ijk} = \delta_k + \gamma Z_j + (\varepsilon_{ij} - \mu_i) , \tag{A2}$$

where δ_k is positively correlated with μ_i by construction.

The negative correlation between the fixed effects δ_k and the error (ε_{ij} -μ_i) will attenuate the estimates of δ_k towards zero. Given the assumption that Z_j is positively correlated with worker ability ($\beta X_i+\alpha_i$), the estimate of γ continues to be biased upwards since worker ability is imbedded in the fixed effect and the associated correlation between Z_j and δ_k biases the coefficient on Z_j upwards. Therefore, the estimates might be improved by directly including X_i in the location fixed effect model specification

$$y_{ijk} = \beta X_i + \delta_k + \gamma Z_j + (\varepsilon_{ij} - \mu_i) . \tag{A3}$$

Again as discussed in the paper, to the extent that δ_k captures worker ability, X_i is also positively correlated with δ_k, and the estimate of β, the coefficient vector for observable human capital conditional on residential fixed effects, will be attenuated relative to the OLS estimates from equation (1) if δ_k is a successful proxy for worker productivity unobservables.

The second limitation associated with these assumptions is that neither commuting costs nor the effect of agglomeration on productivity varies at the individual level (see equation (7)), and in equilibrium these two contributors to wages should be identical. Therefore, if ε_{ijk} and ξ_{ijk} are simply stochastic variations in short-run wages that are unrelated to general features of residential or work location, a model that contains both workplace commuting costs and agglomeration should be unidentified since the two variables should be perfectly collinear (or at least monotonically related allowing for a non-parametric relationship between wages and these variables). Yet empirically, workplace average commute time and our proxies for agglomeration are not perfectly collinear (nor even monotonically related) within metropolitan areas.

One natural explanation for the divergence of agglomeration and commuting time is measurement error in either agglomeration or commuting time. While measurement error in reported commute time might be mitigated by averaging many commute time reports for the same workplace, the effect of agglomeration must be captured by a proxy, such as total employment or employment density. Such proxies likely capture the productivity gains arising from interactions between firms and workers at those firms with considerable error since the ability of firms to share knowledge, labor force, and infrastructure varies with many factors beyond the employment concentration. When agglomeration is captured with measurement error, the relationship between wages and measured agglomeration (Z_j) takes the following form

$$y_{ijk} = \delta_k + \gamma Z_j + (\varepsilon_{ijk} - \gamma\zeta_j), \tag{A4}$$

where $Z_j = \Psi_j + \zeta_j$, and the true level of agglomeration is Ψ_j in work location j, which is orthogonal to the measurement error term ζ_j and perfectly collinear with t_{jk}.[47]

Given that equations (6) and (8) from the paper both hold simultaneously, one can estimate the following model

$$y_{ijk} = \delta_k + \tilde{\eta} t_{jk} + \tilde{\gamma} Z_j + \tilde{\xi}_{ijk} = \delta_k + \eta t_{jk} + 0(\Psi_j + \zeta_j) + \xi_{ijk}. \tag{A5}$$

Under these circumstances, the estimate on t_{jk} will take on its true value since it is orthogonal to the error, while the estimate on agglomeration will be zero since commute time and the true effect of agglomeration are collinear and the agglomeration estimate must be based entirely on the orthogonal measurement error term.[48]

Calculating Bias from Errors in Variables

The problem described above involves bias arising from errors-in-variables with multiple correlated regressors. Given the complexity of this problem, we turn to numeric calculations of the bias in estimated parameters in order to confirm the intuition discussed in the preceding paragraphs. Specifically, we manually calculate the formulas for the omitted variable bias and then calculate the bias implied by assumed values of the variables' variances and covariances based on the observables in our sample. Our calculations will be conducted for four specifications:

$$y_{ij} = X_i + Z_j + \alpha_i + \varepsilon_{ij}, \tag{A6a}$$

$$y_{ijk} = \delta_k + Z_j + (\varepsilon_{ij} - \mu_i), \tag{A6b}$$

[47] Similar to the discussion in footnote 13, Ψ_j and ζ_j can be assumed to be orthogonal without loss of generality by defining ζ_j as the residual arising from a linear projection of the correlated measurement error on Ψ_j, and equilibrium requires that Ψ_j and t_{jk} be collinear conditional on k.

[48] As is shown in the paper, the data are consistent with measurement error in the agglomeration variable, in that the commute time variable captures much more of the variation associated with workplace. If the difference between commute time and agglomeration were associated with measurement error in commuting costs, the estimated coefficient on agglomeration would dominate the coefficient on commute time.

$$y_{ijk} = 0X_i + \delta_k + Z_j + (\varepsilon_{ij} - \mu_i)\,, \tag{A6c}$$

$$y_{ijk} = 0X_i + \delta_k + 0Z_j - t_{jk} + (\varepsilon_{ij} - \mu_i)\,. \tag{A6d}$$

Equation (A6a) is a traditional estimation that is biased by the omission of unobserved productivity or ability variables. Equations (A6b-d) incorporate individual productivity unobservables by including residential location fixed effects, but suffer from bias due to errors-in-variables that arise because residential sorting is driven in part by factors unrelated to total productivity. The "true" coefficients on X_i in (A6c) and (A6d) are zero because total productivity is captured by δ_k, and the "true" coefficient on Z_j in (A6d) is zero because our agglomeration proxy suffers from measurement error and in equilibrium commute time captures the entire effect of agglomeration on wages. The resulting estimates, however, will be non-zero because the variables are correlated with the location fixed effect, which in turn is biased due to the errors-in-variables term μ_i arising from imperfect sorting.[49]

Without loss of generality, all coefficients are initialized to 1 and the impact of a variable on wages is captured by the standard deviation of the variable. Again, without loss of generality, the correlations between X_i and α_i and between $(X_i+\alpha_i)$ and μ_{ik} are assumed to be zero. The models in (A6) are then viewed as reduced form in terms of the individual level regressors where α_i is the residual of unobserved ability that is orthogonal to observed productivity, and μ_i is the residual of individual tastes that are orthogonal to total productivity.[50] For the baseline model,

[49] In principle, the agglomeration estimate may be biased by measurement error, which potentially gives rise to the assumed non-monotonic relationship between agglomeration and commute costs, but our analysis focuses on the bias (in this potentially attenuated estimate) that might arise from the sorting of individuals based on their unobserved productivity. One might examine the bias from measurement error as well. However, we would be uncomfortable making such corrections since measurement error is only one potential explanation for not finding a monotonic relationship between agglomeration and commute time.

[50] See footnote 13 for a precise discussion of the assumption that X_i and α_i are uncorrelated. The assumption that $(X_i+\alpha_i)$ and μ_i are uncorrelated follows a similar logic. This second assumption, however, is only made without loss of generality due to the earlier assumption that individuals can be characterized by the additive sum of $(X_i+\alpha_i)$. If observable and unobservable determinants of productivity have different correlations with unobserved tastes for

the variances of X_i and α_i are initialized to 1. The variances of Z_j and t_{jk} are set to 0.051 and 0.084, respectively. These values were chosen by comparing the standardized estimates of agglomeration from the wage equation (A6c) and the commute time estimate from equation (A6d) relative to the standardized influence of the worker education variables on wages.[51] The correlation between Z_j and t_{jk} is set to 0.74 based on the correlation between workplace agglomeration and average workplace commute time conditional on residential location.[52] The correlation between Z_j and $(X_i + \alpha_i)$ is set to 0.1 in order to allow for sizable bias associated with high productivity individuals sorting into high agglomeration work locations. Next, the variance of the residential location taste unobservable is set to 3 in order to match the observed attenuation of the estimates on the human capital variables of approximately 25% when residential fixed effects are included in the model (A6c).[53] Finally, the correlation between $(X_i + \alpha_i)$ and t_{jk} is set to zero initially, and this correlation is investigated later in this appendix.

Table A1 presents the expectation for parameter estimates or the sum of the true value plus the bias using standard omitted variable calculations.[54] The first panel presents the baseline

location, then X_i and α_i would not enter the fixed effect in a reduced form model with the same weights as they enter the wage equation.

[51] Specifically, an education index is constructed using the estimated coefficients on the educational attainment dummy variables. The standardized coefficients on employment density in our fixed effects model (see Table 3) is approximately 0.225 times the standard deviation of the education index, and the standardized coefficient on workplace average commute time is approximately 0.290 times the education index standard deviation. These standardized effects are based on conditioning out other individual controls and metropolitan area fixed effects.

[52] The non-unitary correlation between agglomeration and commuting costs when combined with the initialization of the agglomeration coefficient to zero is consistent with measurement error in agglomeration, but not in commuting costs. The empirical correlation between employment density and average workplace commute time is also conditional on metropolitan area fixed effects and all controls other than the human capital variables.

[53] The attenuation of the coefficients for educational attainment dummy variables is between 22 and 23 percent in the initial model that controls for census tract fixed effects, and attenuation increases to 24-26 percent with block group fixed effects and to 26-29 percent with housing submarket by census tract fixed effects.

[54] The expected value of parameter estimates can be calculated using the underlying model rather than the more typical least squares calculations, which require a specification for the fixed effects model such as the inclusion of residential location dummy variables. Rather, the expected value of wages conditional on the fixed effect model is

$$E[y_{ijk} \mid \delta_k, Z_j] = \delta_k + Z_j - E[\mu_{ik} \mid \delta_k, Z_j]$$

and the expectation of the unobservable can be expressed as a linear function of the fixed effect and an orthogonal regressor if expectations are assumed to be a linear function of conditioning variables expectations of estimates given the variances and correlations described above, and the following panels present expectations after changing one of the variance-covariance terms. The baseline results show that the OLS estimate in column 1 is biased above the true value of 1. The bias on the agglomeration variable is actually increased by replacing observable human capital measures with residential location fixed effects (column 2). This increase arises from the high variance assigned the taste unobservable, and bias is decreased between equations (A6a) and (A6b) in models where that variance is less than 2.0. Nonetheless, column 3 illustrates that the bias is reduced by the inclusion of residential fixed effects into a model that controls for observable productivity or human capital (A6c). The inclusion of commute time in the fourth and final column (A6d) dramatically reduces the estimates on the agglomeration variable. Notably, the coefficient on the agglomeration variable after controlling for commutes (column 4) is larger than the bias on the agglomeration estimates after controlling for residential location fixed effects and observed human capital (column 3) and so provides an upper bound on the bias from imperfect sorting. Finally, looking at the second row of panel 1, the attenuation in the coefficient estimate on human capital is about 0.25 as calibrated to be consistent with attenuation in our empirical models. This attenuation decreases monotonically with the variance of the taste unobservable.

While the magnitude of the bias changes with the variance and covariance terms, the basic pattern of results remains the same. Decreasing the relative contribution of agglomeration

---

$E[\mu_{ik} \mid \delta_k, Z_j] = \alpha_0 + \alpha_1 \delta_k + \alpha_2 (Z_j - E[Z_j \mid \delta_k]) + \alpha_2 E[Z_j \mid \delta_k] = (\alpha_0 + \alpha_2 \gamma_0) + (\alpha_1 + \alpha_2 \gamma_1)\delta_k + \alpha_2 (Z_j - E[Z_j \mid \delta_k])$
where α_2 captures the bias in the coefficient on Z_j, but this bias involves a conditional expectation, $E[Z_j \mid \delta_k] = \gamma_0 + \gamma_1 \delta_k$. In order to calculate the bias in terms of unconditional moments, we recognize that γ_1=Cov[Z_j, δ_k]/Var[δ_k] and ($\alpha_1 + \alpha_2 \gamma_1$)= Cov[$\mu_{ik}$, δ_k]/Var[δ_k], and then reversing the process yields an equivalent coefficient on Z_j in a model where the other regressor is orthogonal. The resulting two equations can be solved for the bias, and the results are identical to the results of the least squares omitted variable calculation in the case where one actually observes the true fixed effect and can include it as a regressor.

to wages (panel 2), increasing the contribution of unobserved ability (panel 3), or increasing the correlation between individual productivity (both observed and unobserved) and agglomeration (panel 4) all increase the bias in agglomeration estimates, but the bias is still reduced by including fixed effects in a model with human capital controls (column 3), and the agglomeration coefficient in the model with commute time (column 4) still provides an upper bound to the bias on agglomeration estimate in column 3.[55] Finally, decreasing the correlation between agglomeration and commute time to zero (panel 5), which must be positively related in equilibrium, leads to an agglomeration parameter in column 4 that is the same as the bias in column 3 and so the column 4 estimate still provides an upper bound for bias.[56] While the expectation calculations are based on one observable measure of productivity, we have repeated these calculations with multiple measures, and the results of those calculations are very similar regardless of the correlations assumed between the observable productivity variables.

The one exception to these findings arises from a correlation between productivity and commute time. A positive correlation between commute time and an individuals' productivity decreases the expectation for the coefficient on the agglomeration variable in column four, and so this expectation may no longer provide an upper bound for the bias in the agglomeration estimate from the model in column 3. This finding is not surprising. As discussed earlier, a key threat to the validity of our second test for bias from unobserved ability, where we ask whether agglomeration effects on wages can be explained by or compensated away by commuting costs, is the sorting of households across commute times based on ability.

[55] Note that we also increase the variance of the taste unobservable when we increase the variance of unobserved ability in panel 3 in order to recalibrate the attenuation on the human capital estimate.
[56] The zero correlation is an extreme case. Since commute time and the true productivity effect of agglomeration are collinear, zero correlation implies sufficient measurement error to render our agglomeration proxy meaningless, and so whenever the agglomeration proxy is informative, estimates from the commute time model should provide an upper bound on bias from sorting.

Table A2 repeats the calculations of the expected value of estimates using correlations for productivity (both observed and unobserved) with agglomeration and commute time drawn from the data. Specifically, we calculate the correlation between our education index and both agglomeration and commute time and use those as the correlation with both the observable and the unobservable component of productivity. After conditioning on metropolitan area and other individual observables, the estimate of the correlation between our measure of observable productivity, education level, and agglomeration is approximately 0.040, and the estimate of the correlation between observable productivity and average workplace commute time is 0.060. Panel 1 shows the expectations based on these estimates, and the agglomeration estimate in column 4 does not provide an upper bound on the bias in column 3. The correlation between innate productivity and commute time must fall below 0.029 for column 4 to provide an upper bound (panel 2). This phenomenon arises in part because of the very small assumed correlation between worker innate productivity and agglomeration and so the failure of column 4 is to a large extent associated with situations where there is little bias in the agglomeration estimates. Panel 3 shows that column 4 provides an upper bound if that correlation between innate productivity and agglomeration rises above 0.081.

These calculations indicate that column 4 provides an upper bound for the bias in column 3 when the expectation of the commute time estimate is equal to or below the true value. In the Table A1 calculations, the expectation of the commute time estimate is always considerably less than one, and column 4 provides an upper bound with substantial clearance. In panel 1 of Table A2, the estimate for commute time is biased upwards by 18 percent and column 4 does not provide an upper bound, while in panels 2 and 3 we chose the covariances so that the expected value for commute time is 1.0, and this assumption leads to estimates in column 4 that exactly capture the bias in the agglomeration estimates from column 3. This finding is consistent with the earlier intuition that agglomeration effects should be completely compensated for by commuting costs, but that too high a coefficient estimate on commute time suggests bias because households sort across commutes and/or workplaces based on unobserved productivity.

Finally, across both tables, these simulations indicate a 25 percent reduction in bias from the inclusion of residential fixed effects that attenuate the human capital coefficients by 25 percent for a wide array of parameter values except for panels 3 and 4 in Table A1. In panel 3, increasing the variance of unobserved ability leads to a smaller reduction in bias of 23.3 percent compared to the panel 1 reduction of 24.8, but this change is primarily associated with the change in attenuation of the human capital estimate from 25.6 to 24.4 percent because if the variance of the neighborhood preference parameter is reduced to 6.1 so that attenuation in the human capital estimate remains identical across panels the reduction in bias is 24.6, very close to the panel 1 value. In panel 4, doubling the correlation between productivity and agglomeration decreases the reduction in bias slightly and increases the attenuation in the human capital estimate. Even when the variance of the preference parameter is increased to 3.3 to exactly match the attenuation of the human capital estimate in panel 1, a doubling of this correlation only reduces the reduction in bias from 24.8 in panel 1 to 22.3 percent. The changes in Table A2 have little effect on the percent reduction in bias. In summary, the reduction is bias is substantial for a large range of parameter values, but quite sensitive to the amount of attenuation in the human capital estimate. A 1.2 percentage point reduction in the attenuation leads to a 1.3 percentage point decrease in the reduction in bias from including residential fixed effects.

Table 1: Variable Names, Means, and Standard Deviations		
Variable Name	Non-College	College Graduates
Dependent Variable		
Average hourly wage	20.103 (30.828)	35.987 (55.428)
Workplace Controls		
Total Residential PUMA employment in 100,000’s	0.488 (0.575)	0.641 (0.759)
PUMA Employment density in 1000’s/square KM	2.646 (11.004)	4.772 (15.306)
Share of college educated workers in PUMA	0.358 (0.094)	0.405 (0.101)
Average commute time to PUMA in minutes	26.573 (6.629)	28.195 (7.787)
Metropolitan Area Controls		
Percent college educated in MSA and occupation	0.026 (0.035)	0.056 (0.045)
Percent college educated in MSA and industry	0.033 (0.028)	0.051 (0.035)
Individual Worker Controls		
Age of worker	42.580 (7.980)	43.024 (8.076)
Non-Hispanic white worker	0.672 (0.470)	0.813 (0.390)
African-American worker	0.126 (0.332)	0.058 (0.233)
Hispanic worker	0.159 (0.365)	0.043 (0.204)
Asian and Pacific Islander worker	0.042 (0.200)	0.084 (0.278)
High school degree	0.346 (0.476)	
Associates degree	0.488 (0.500)	
Four year college degree		0.599 (0.490)
Master degree		0.255 (0.436)
Degree beyond Masters		0.146 (0.353)
Worker single	0.285 (0.452)	0.230 (0.421)
Presence of own children in household	0.474 (0.499)	0.502 (0.500)
Born in the United States	0.800 (0.400)	0.826 (0.379)
Years in residence if not born in U.S.	18.574 (10.809)	17.432 (11.669)
Quality of spoken English	0.164 (0.370)	0.168 (0.374)
Sample size	141,5176	92,7916

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.

Table 2: Baseline Model of Agglomeration Economies for Logarithm of the Wage Rate		
Independent Variables	Total Employment	Density
Total employment in 100,000’s	0.0544 (7.80)	
Employment density in 1000’s per square KM		0.0024 (8.40)
Percent college educated in MSA and occupation	0.8762 (6.36)	0.9237 (6.56)
Percent college educated in MSA and industry	1.7127 (10.27)	1.7520 (10.69)
Age of worker	0.0333 (46.19)	0.0333 (45.99)
Age of worker squared divided by 100	-0.0369 (-46.51)	-0.0369 (-46.32)
Non-Hispanic white worker	0.1376 (28.00)	0.1368 (27.71)
African-American worker	-0.0064 (-1.33)	-0.0059 (-1.23)
Hispanic worker	-0.0152 (-3.04)	-0.0156 (-3.12)
Asian and Pacific Islander worker	0.0359 (5.73)	0.0359 (5.68)
High school degree	0.1380 (57.50)	0.1383 (57.90)
Associates degree	0.2241 (73.75)	0.2250 (74.84)
Four year college degree	0.4219 (111.51)	0.4244 (113.05)
Master degree	0.5429 (104.14)	0.5463 (107.67)
Degree beyond Masters	0.6606 (121.29)	0.6645 (125.10)
Worker single	-0.1350 (-107.90)	-0.1346 (-107.07)
Presence of own children in household	0.0720 (46.93)	0.0717 (45.93)
Born in the United States	-0.0563 (-20.81)	-0.0564 (-20.75)
Years in residence if not born in U.S.	0.0087 (59.42)	0.0087 (59.31)
Quality of spoken English	0.0135 (4.88)	0.0135 (4.87)
R-square	0.2905	0.2898

Note: The dependent variable for all regressions is the logarithm of the estimated hourly wages, which is calculated as annual labor market earnings divided by the product of number of weeks worked and average hours worked per week. The key variable of interest is either the total number of full time workers in an individual’s workplace based on residential PUMA or the density of full time workers in the workplace where full-time work is defined as worked an average of at least 35 hours per week. The sample of 2,343,092 observations contains male full-time workers aged 30 to 59 in the selected metropolitan areas. The models include metropolitan area, 15 industry, and 20 occupation fixed effects, but those estimates are suppressed. T-Statistics based on standard errors clustered at the workplace are shown in parentheses.

Table 3: Agglomeration Wage Models without and with Location Controls						
Variables	Total Employment			Density		
	OLS	Fixed Effects	Commute Time	OLS	Fixed Effects	Commute Time
Baseline Model Specification						
Employment	0.0544 (7.80)	0.0508 (9.79)	0.0082 (2.62)			
Density				0.0024 (8.40)	0.0026 (7.98)	0.0004 (3.83)
Commute Time			0.0067 (20.33)			0.0069 (28.51)
R-Square	0.2905	0.3340	0.3347	0.2898	0.3338	0.3347
Logarithm of Employment, Density, and Commute Time						
Employment	0.0516 (22.02)	0.0432 (17.78)	0.0137 (7.31)			
Density				0.0181 (11.88)	0.0202 (13.55)	0.0047 (5.65)
Commute Time			0.1855 (16.72)			0.1972 (17.11)
R-Square	0.2912	0.3340	0.3346	0.2902	0.3340	0.3346
No Individual Level Covariates						
Employment	0.0544 (7.01)	0.0575 (9.33)	0.0101 (2.92)			
Density				0.0022 (8.15)	0.0029 (7.51)	0.0003(3.10)
Commute Time			0.0075 (20.12)			0.0079 (29.43)
R-Square	0.1997	0.2895	0.2904	0.1987	0.2892	0.2904
Sample of Single Men						
Employment	0.0409 (7.45)	0.0368 (8.16)	0.0004 (0.14)			
Density				0.0018 (7.93)	0.0018 (7.04)	-0.0001 (-0.88)
Commute Time			0.0062 (16.02)			0.0064 (19.85)
R-Square	0.2427	0.3078	0.3084	0.2422	0.3076	0.3084
Observationally Equivalent Cells						
Employment	0.0569 (8.40)	0.0533 (8.98)	0.0068 (1.81)			
Density				0.0025 (8.17)	0.0028 (7.08)	0.0003 (2.16)
Commute Time			0.0071 (18.35)			0.0073 (20.55)
R-Square	0.2930	0.3563	0.3570	0.2923	0.3560	0.3570

Note: The OLS columns in panel 1 contain the results from Table 2, the fixed effect columns contain the results for the same model where metropolitan fixed effects are replaced by census tract of residence fixed effects, and the commute time columns contain the results for the census tract fixed effect model after the inclusion of the average commute time for the individual's workplace at the residential PUMA level. Panel 2 presents estimates controlling for the logarithm of total employment or employment density, as well as the logarithm of average commute time for the last model. Panel 3 presents estimates for a specification where all individual worker covariates (as listed in Table 1) are excluded, panel 4 present estimates for a sample of single men, and panel 5 presents estimates based on a model that controls for worker cell by census tract fixed effects. The first three and the last models use the same sample of 2,343,092 observations, while the fourth model uses the subsample of single men with 617,144 observations. T-Statistics based on standard errors clustered at the workplace are shown in parentheses.

Table 4: Agglomeration Wage Models Instrumenting for Commute Time as Share of Work Day						
Variables	Total Employment			Density		
	Commute Time IV Estimation	Commute Coefficient 1.5	Commute Coefficient 1.0	Commute Time IV Estimation	Commute Coefficient 1.5	Commute Coefficient 1.0
Employment	0.0082 (2.62)	0.0149 (6.64)	0.0268 (8.82)			
Density				0.0004 (3.83)	0.0008 (8.13)	0.0014 (8.47)
Commute Time	1.7766 (20.33)	1.5000	1.0000	1.8246 (28.51)	1.5000	1.0000
R-Square	0.3347	0.3287	0.3300	0.3347	0.3287	0.3300

Note: The first and fourth columns present two-stage least squares estimates for the census tract of residence fixed effect agglomerations models controlling for an individual's total commute time (both ways) as a share of their entire work day (average hours worked per week divided by five plus the total commute time) using the average commute time for the workplace based on residential PUMA (the same control variable used in Table 3) as an instrument. The next two columns present estimates based on predicted commute time share, but restricting the coefficient on commute time share to 1.5 and 1.0, respectively. T-Statistics based on standard errors clustered at the workplace are shown in parentheses. The sample size is 2,343,092.

Table 5: Employment Density Models with Alternative Workplace Definitions			
Variables	OLS	Fixed Effects	Commute Time
Workplace PUMA			
Density	0.0063 (4.31)	0.0074 (4.66)	0.0002 (0.43)
Standardized Density	0.0246	0.0289	0.0008
Commute Time			0.0075 (22.81)
R-Square	0.2892	0.3333	0.3341
Residential PUMA			
Density	0.0024 (8.40)	0.0026 (7.98)	0.0004 (3.83)
Standardized Density	0.0310	0.0336	0.0052
Commute Time			0.0069 (28.51)
R-Square	0.2898	0.3338	0.3347
Zip Code Tabulation Area			
Density	0.0013 (4.44)	0.0012 (4.12)	0.0003 (3.65)
Standardized Density	0.0297	0.0274	0.0069
Commute Time			0.0077 (30.72)
R-Square	0.2891	0.3340	0.3364

Note: The workplace geography for each panel is used to calculate employment density in and average commute time to a workplace for the models presented in that panel. The estimates in panel 2 contain the results from Table 3 where workplace is defined based on residential Public Use Microdata Areas (PUMAs). Panel 1 defines workplace using the larger workplace PUMAs, and panel 3 using the five-digit census defined zip code tabulation areas. Estimates on employment density are scaled or standardized using the within metropolitan area standard deviation of that variable for the specific geography. The standard deviations for employment density are 3.9088, 12.9226, and 22.8446 for the workplace PUMA, residential PUMA and Zip Code Tabulation Area, respectively. All fixed effect models (column two) include census tract of residence fixed effects. The models include the standard covariates shown in Table 1, and estimates are based on the full sample of 2,343,092 observations for panels 1 and 2 and on 2,132,986 observations for panel 3. T-Statistics are shown in parentheses and based on standard errors clustered at the workplace geography used in each panel.

Table 6: Employment Density Models with Alternative Residential Neighborhood Definitions			
Variables	OLS	Fixed Effects	Commute Time
Residential PUMA			
Density	0.0024 (8.40)	0.0028 (8.32)	0.0003 (2.97)
Commute Time			0.0078 (28.83)
R-Square	0.2898	0.3036	0.3048
Zip Code Tabulation Area			
Density		0.0027 (7.90)	0.0004 (3.68)
Commute Time			0.0071 (28.00)
R-Square		0.3150	0.3160
Census Tract			
Density		0.0026 (7.98)	0.0004 (3.83)
Commute Time			0.0069 (28.51)
R-Square		0.3338	0.3347
Census Block Group			
Density		0.0026(7.93)	0.0004(3.96)
Commute Time			0.0068(28.44)
R-Square		0.3600	0.3609

Note: The residential neighborhood geography for each panel is used to define the residential location fixed effects. The estimates in panel 3 contain the results from Table 3 where fixed effects are defined using census tracts. Panel 1 defines the fixed effects using residential PUMAs, panel 2 using the five-digit census defined zip code tabulation areas, and panel 4 using census block groups. All models define employment density and average commute time based on workplace at the residential PUMA level. The models include the standard covariates shown in Table 1, and use the full sample of 2,343,092 observations. T-Statistics are shown in parentheses and based on standard errors clustered at the workplace.

Table 7: Employment Density Models with Alternative Neighborhood Fixed Effects			
Variables	OLS	Fixed Effects	Commute Time
Census Tract Fixed Effects			
Density	0.0024 (8.40)	0.0026 (7.98)	0.0004 (3.83)
Commute Time			0.0069 (28.51)
R-Square	0.2898	0.3338	0.3347
Census Tract by Tenure in Residence Fixed Effects			
Density		0.0025 (8.08)	0.0004 (4.56)
Commute Time			0.0066 (27.47)
R-Square		0.3701	0.3709
Census Tract by Housing Stock Fixed Effects			
Density		0.0025 (8.27)	0.0004 (4.36)
Commute Time			0.0066 (27.47)
R-Square		0.3854	0.3862

Note: All models use workplace agglomeration and commute time at the residential PUMA level. The models in panel 1 control for census tract fixed effects. The models in panel 2 control for tenure based fixed effects that include a unique fixed effect for each of four tenure categories in each census tract. The models in panel 3 control for housing stock fixed effects that include a unique fixed effect for each housing stock category in each census tract. The four tenure categories are renter, owner in residence less than one year, owner in residence between one and five years, and owner in residence more than five years. The seven housing stock categories are mobile home, multifamily 1 bedroom or less, multifamily 2 bedroom, multifamily 3 bedroom or more, single family 2 or less bedrooms, single family 3 bedrooms, and single family 4 or more bedrooms. The models include the standard covariates shown in Table 1, and T-Statistics are shown in parentheses and based on standard errors clustered at the workplace. Sample size: 2,343,092.

Table 8: Employment Density Models for Different Regions of the Country			
Variables	OLS	Fixed Effects	Commute Time
Full Sample			
Raw Density	0.0024 (8.40)	0.0026 (7.98)	0.0004 (3.83)
Standardized Density	0.0310	0.0336	0.0052
Commute Time			0.0069 (28.51)
R-Square	0.2898	0.3338	0.3347
Sample Size	2,343,092		
Northeast			
Raw Density	0.0022 (9.65)	0.0023 ((10.18)	0.0004 (3.34)
Standardized Density	0.0523	0.0547	0.0095
Commute Time			0.0066 (19.4)
R-Square	0.2923	0.3352	0.3365
Sample Size	569,806		
Midwest			
Raw Density	0.0037 (12.64)	0.0038 (10.65)	0.0004 (0.9)
Standardized Density	0.0258	0.0265	0.0028
Commute Time			0.0067 (10.68)
R-Square	0.2644	0.3079	0.3085
Sample Size	527,781		
South			
Raw Density	0.0052 (6.4)	0.0046 (5.87)	0.0007 (1.36)
Standardized Density	0.0167	0.0148	0.0023
Commute Time			0.0067 (10.27)
R-Square	0.3065	0.3485	0.3492
Sample Size	637,023		
West			
Raw Density	0.0032 (2.58)	0.0047 (4.74)	-0.00005(-0.12)
Standardized Density	0.0116	0.0171	-0.0002
Commute Time			0.0069 (13.89)
R-Square	0.2905	0.3356	0.3362
Sample Size	608,482		

Note: All models use workplace agglomeration and commute time at the residential PUMA level and fixed effect models control for census tract fixed effects. The estimates in panel 1 are for the full sample and the estimates in panels 2-5 are for the subsample of residents residing in the Northeast, Midwest, South, and West census regions. The standardized density coefficients are based on the within metropolitan area standard deviation of the employment density variable measured at the workplace PUMA. The standard deviations are 12.9226, 23.7880, 6.9753, 3.2174, and 3.6324 for the full sample, Northeast, Midwest, South, and West, respectively. The models include the standard covariates shown in Table 1, and T-Statistics are shown in parentheses and based on standard errors clustered at the workplace.

Table 9: Employment Density Models for Subgroups			
Variables	OLS	Fixed Effects	Commute Time
No Four Year College Degree			
Raw Density	0.0017 (6.91)	0.0023 (7.10)	-0.00003 (-0.35)
Standardized Density	0.0187	0.0253	-0.0003
Commute Time			0.0074 (26.41)
R-Square	0.2156	0.2633	0.2646
Sample Size	1,415,176		
Four Year College Degree			
Raw Density	0.0030 (8.86)	0.0028 (8.63)	0.0008 (5.27)
Standardized Density	0.0459	0.0429	0.0122
Commute Time			0.0064 (18.46)
R-Square	0.1750	0.2514	0.2522
Sample Size	927,916		
Non-Hispanic White			
Raw Density	0.0030 (8.61)	0.0030 (8.21)	0.0003 (2.60)
Standardized Density	0.0369	0.0369	0.0037
Commute Time			0.0081 (29.99)
R-Square	0.2473	0.2987	0.2999
Sample Size	1,705,058		
Minority			
Raw Density	0.0011 (7.39)	0.0017 (8.31)	0.0007 (4.44)
Standardized Density	0.0158	0.0245	0.0101
Commute Time			0.0034 (8.06)
R-Square	0.2859	0.3507	0.3510
Sample Size	638,034		
Automobile Commuter			
Raw Density	0.0033 (7.34)	0.0032 (7.41)	0.0004 (3.61)
Standardized Density	0.0224	0.0218	0.0027
Commute Time			0.0072 (27.19)
R-Square	0.2831	0.3281	0.3290
,,Sample Size	2,073,487		
Mass-Transit Commuter			
Raw Density	0.0022 (7.03)	0.0016 (6.68)	0.0001 (0.54)
Standardized Density	0.0735	0.0534	0.0033
Commute Time			0.0069 (5.19)
R-Square	0.4243	0.5360	0.5367
Sample Size	144,917		

Note: All models use workplace agglomeration and commute time at the residential PUMA level and fixed effect models control for census tract fixed effects. The estimates in panels 1 and 2 are for the subsamples without and with a four year college degree, panels 3 and 4 are for the non-Hispanic white and minority subsamples, and panels 5 and 6 are for automobile and mass-transit commuter subsamples. The standardized density coefficients are based on the within metropolitan area standard deviation of the employment density variable for each sample measured at the workplace PUMA. The standard deviations are 11.0040, 15.3060, 12.3109, 14.3943, 6.8012, and 33.3989 in order of the panels. The models include the standard covariates shown in Table 1, and T-Statistics are shown in parentheses and based on standard errors clustered at the workplace.

Table 10: Employment Density and Workplace Human Capital Models			
Variables	OLS	Fixed Effects	Commute Time
Baseline Model Specification			
Density	0.0015 (11.64)	0.0022 (8.55)	0.0004 (3.86)
Share Workers with College	0.3593 (17.68)	0.1512 (10.94)	0.0472 (3.30)
Commute Time			0.0066 (24.92)
R-Square	0.2913	0.3340	0.3347
No Individual Level Covariates			
Density	0.0010 (8.28)	0.0024 (8.05)	0.0003(3.18)
Share Workers with College	0.4793 (19.66)	0.1990 (13.00)	0.0833(5.29)
Commute Time			0.0073 (24.81)
R-Square	0.2013	0.2895	0.2904
Sample of Single Men			
Density	0.0012 (9.28)	0.0015 (7.17)	-0.0001 (-0.97)
Share Workers with College	0.2596 (12.19)	0.1359(8.55)	0.0401 (2.54)
Commute Time			0.0061 (18.03)
R-Square	0.2431	0.3077	0.3084
Observationally Equivalent Cells			
Density	0.0017 (11.05)	0.0024 (7.60)	0.0003 (2.20)
Share Workers with College	0.3534(18.74)	0.1719 (9.55)	0.0597 (3.30)
Commute Time			0.0068 (18.01)
R-Square	0.2937	0.3563	0.3570

Note: Panel 1 presents estimates from the baseline specification presented in Table 3 extended to include a control for the share of workers with a college degree at the workplace. Panel 2 presents estimates for a specification where all individual worker covariates (as listed in Table 1) are excluded, panel 3 presents estimates for a sample of single men, and panel 4 presents estimates based on a model that controls for worker cell by census tract fixed effects. The first two and the last models use the same sample of 2,343,092 observations while the third model uses the subsample of single men with 617,144 observations. T-Statistics based on standard errors clustered at the workplace are shown in parentheses.

Table A1: Calculation of the Expectation of Parameter Estimates				
Parameters	Ordinary Least Squares	Residential Fixed Effect	Residential Fixed Effect plus Observables	Residential Fixed Effect plus Observables and Commutes
Baseline				
Agglomeration	1.447	1.536	1.336	0.755
Human Capital	0.990		0.744	0.737
Decrease Variance of Agglomeration from 0.051 to 0.010				
Agglomeration	2.010	2.210	1.759	1.705
Human Capital	0.990		0.744	0.737
Increase Variance of Unobserved Ability from 1.0 to 2.0				
Agglomeration	1.633	1.736	1.485	1.088
Human Capital	0.986		0.756	0.744
Increase Correlation between Agglomeration and Human Capital from 0.1 to 0.2				
Agglomeration	1.923	2.098	1.699	1.651
Human Capital	0.958		0.726	0.694
Decrease Correlation between Agglomeration and Commute time from 0.74 to 0.0				
Agglomeration	1.447	1.536	1.336	0.336
Human Capital	0.990		0.744	0.744

Note: The cells contain the true value of the parameter plus the calculated bias based on the models specified in equations (A6a-d). The baseline calculations are based on a variance of X_i and α_i of 1, a variance of Z_j of 0.051, a variance of t_{jk} of 0.084, a correlation between Z_j and (X_i + α_i) of 0.1, a correlation between Z_j and t_{jk} of 0.74, and a correlation between (X_i + α_i) and t_{jk} of 0. All baseline values are preserved in following panels except for the specific variance or correlation being modified in the panel. In panel 3, the variance of the residential preference parameter increases from 3.0 to 6.5 in order to keep the attenuation of human capital variables in model 3 approximately constant.

Table A2: Calculation of the Expectation of Parameter Estimates				
Parameters	Ordinary Least Squares	Residential Fixed Effect	Residential Fixed Effect plus Observables	Residential Fixed Effect plus Observables and Commutes
Correlations of Agglomeration with Human Capital at 0.040 and with Commute Time at 0.060				
Agglomeration	1.177	1.213	1.133	-0.032
Human Capital	0.998		0.749	0.748
Commute Time				1.175
Decrease Correlation between Agglomeration and Commute Time to 0.029				
Agglomeration	1.177	1.213	1.133	0.136
Human Capital	0.998		0.749	0.749
Commute Time				0.997
Increase Correlation between Agglomeration and Human Capital to 0.081				
Agglomeration	1.361	1.433	1.271	0.271
Human Capital	0.993		0.746	0.746
Commute Time				1.000

Note: The cells contain the true value of the parameter plus the calculated bias based on the models specified in equations (A6a-d). The panel 1 calculations are based on a variance of X_i and α_i of 1, a variance of Z_j of 0.051, a variance of t_{jk} of 0.084, a correlation between Z_j and $(X_i + \alpha_i)$ of 0.040, a correlation between Z_j and t_{jk} of 0.74, and a correlation between $(X_i + \alpha_i)$ and t_{jk} of 0.060. All panel 1 values are preserved in following panels except for the specific correlation being modified in the panel.