The Urban Density Premium across Establishments

R. Jason Faberman, Federal Reserve Bank of Chicago
Matthew Freedman, Department of Economics, Cornell University*

April 2012

Abstract
We examine the earnings premium associated with urban density from the perspective of establishments. We do so in an analogous manner to studies on the urban density premium that focused on individuals. We establish several facts that are comparable to the density premium afforded workers: establishments reap an overall return to density just over 7 percent after various controls and the highest-earnings establishments exhibit a density premium that is about 1.5 times greater than the lowest-earnings establishments. We find evidence against theories of firm learning or firm selection through exit: the density premium does not rise with establishment age and there is no difference in exit rates across the earnings distributions of high- and low-density metropolitan areas. We find some evidence of sorting, though towards *lower-density* metropolitan areas, which we interpret as consistent with a notion of dense cities as "nursery cities."

Keywords: density premium, urban agglomeration, firm sorting and selection, establishment entry and exit
JEL Codes: D22, R12, R30

** Correspondence: Faberman: jfaberman@frbchi.org, Freedman: Freedman@cornell.edu.* *We thank Nate Baum-Snow, Laurent Gobillon, and Ronni Pavan, as well as seminar participants at the Urban Economics Association Meetings and the Conference for Urban and Regional Economics, for helpful comments. The views in this paper are our own and do not reflect the views or the Federal Reserve Bank of Chicago, the Federal Reserve System, or the views of its staff members. In addition, research in this paper was conducted while the authors were Special Sworn Status researchers of the U.S. Census Bureau at the New York and Chicago Census Research Data Centers. Research results and conclusions expressed are those of the authors and do not necessarily reflect the views of the Census Bureau. This paper has been screened to insure that no confidential data are revealed.*

1. Introduction

For years, urban economists have studied why observationally similar workers earn more in more densely populated locations. Studies have consistently found an elasticity of earnings with respect to urban density between 3 and 10 percent. This elasticity is often robust to controlling for a variety of factors, including the role of migration and the sorting of skilled workers across cities, the role of the returns to worker experience at a particular location, and the role of labor search and matching frictions.[1]

The existing research on the earnings density premium has focused primarily on the behavior of workers. The fact that the earnings premium persists after controlling for the above factors suggests that a sizable portion could stem from benefits realized on the firm side of the labor market. While there is a growing literature on how worker characteristics and worker behavior relate to the density premium, little is known about the analogous relations for firms.[2] In this paper, we document the establishment-level behavior that gives rise to the empirical relation between earnings and the density of a metropolitan area. Using a rich source of longitudinal establishment micro data, we proceed in a way that closely mirrors previous research on workers, focusing on the average earnings at a particular establishment. We also argue that an establishment's average earnings reflects differences in productivity, so one can interpret the establishment-based premium as capturing the potential productive benefits an establishment receives from locating in a dense city. In this regard, we also build on the handful of studies that relate firm productivity to urban agglomeration (e.g., Syverson, 2004; Combes et al., 2009; Lehmer and Möeller, 2010). Finally, we relate our evidence to theories of the urban density premium and worker and firm behavior. Specifically, we focus on three theories. The first is firm

1 See Glaeser (1999), Glaeser and Mare (2001), Freedman (2008), Bacolod, Blum, and Strange (2009), Glaeser and Resseger (2010), Baum-Snow and Pavan (2010a, 2010b), and de la Roca and Puga (2010), among others.
2 Notable exceptions include Ciccone and Hall (1996), Duranton and Puga (2001) and Moretti (2004), though these studies focus more specifically on the role of agglomeration economies and spillovers instead of the earnings premium.

learning. Glaeser and Mare (2001) and Baum-Snow and Pavan (2011) show that worker's wage-tenure profiles are steeper in larger cities, and attribute this to faster human capital accumulation, or learning. We examine whether similar "learning" effects over time (e.g., through accumulated knowledge spillovers, or learning about product demand) are present for firms. The second is firm selection through exit. Syverson (2004) shows that, within the concrete industry, a higher density of product demand leads to greater competition, which in turn leads to more exits among low-productivity concrete plants and higher mean productivity through a greater lower-truncation of the productivity distribution. The third is firm sorting. The endogenous sorting of skilled workers into agglomerated areas is a common concern in the literature. For example, Combes et al. (2010) find that such sorting accounts for about 35 percent of their estimated density premium. We examine how much sorting by firms across metropolitan areas accounts for our estimates.

We begin with some stylized facts about the relationship between average establishment earnings and population density across U.S. metropolitan areas. As Figure 1 shows, the elasticity of metropolitan-area earnings with respect to population density is about 8 percent. At the establishment level, we estimate an elasticity of average earnings with respect to density of about 10 percent, which falls to 7.4 percent when controls for the share of the population with a college degree and establishment characteristics are added. We obtain roughly similar estimates by industry, establishment size, and other characteristics. We also find that instrumenting to control for the potential endogeneity of density (what Combes et al., 2010, refer to as the "endogenous quantity of labor") does little to affect our estimates. We also find that the estimated density premium is increasing in average establishment earnings. The spread between establishments in the lowest earnings quintile and those in the top earnings quintile is 6.4 percentage points unconditionally and 3.5 percentage points after adding controls. The finding is analogous to findings of steeper wage-tenure profile for workers in larger cities by, for example, Glaeser and Mare (2001), Gould (2007), and Glaeser and Resseger (2010). The finding is also consistent with Combes et al. (2009), who find that the returns to TFP of locating in a large city are 14 percent for the top quartile of firms and 5 percent for the bottom quartile. Thus, like Combes et al., we find evidence of what we refer to as "productivity-biased" returns to agglomeration.

We then turn to examining evidence of firm learning, sorting, and selection. For firm learning, we examine whether the estimated density premium rises with establishment age. Since establishments rarely relocate across cities, establishment age and city tenure are roughly equal. Thus, if the density premium increases with establishment age, it would be evidence of accumulated returns to urban density, or "learning," by establishments. We find strong evidence *against* learning on the part of establishments. Establishments realize their returns to density at entry, and the estimated premium is essentially constant after that. This holds after adding controls for the college share and establishment characteristics, and within various subgroups of the data. The finding contrasts with previous research on workers that finds workers in denser cities have relatively steeper earnings profiles with respect to city tenure (e.g., Glaeser and Mare, 2001; Baum-Snow and Pavan, 2010a; de la Roca and Puga, 2011).

Theories of firm selection imply that locations with a greater selection effect should have higher productivity through the lower-truncation of their productivity distribution. This lower-truncation occurs through the exit of low-productivity firms. One concern is that the estimated density premium is partly due to a greater selection effect in more dense areas. We test for evidence of a greater selection effect two ways. First, we estimate exit rates as a function of the within-city earnings distribution for high-density and low-density cities. While we find higher exit rate at the lower tail of the earnings distribution, we find no *differences* in these exit rates with respect to density. Second, track the behavior of the earnings distribution of a cohort of entrants for five years in high-density and low-density cities. [TO BE COMPLETED.]

Finally, a common concern for estimating returns to urban agglomeration is the sorting of high-skill workers, or in our case high-productivity establishments, into denser areas. Like selection, such sorting could partly account for our estimated density premium. Workers sort across cities through migration, but establishments can sort along two margins. The first is through entry, since establishments can potentially choose *ex ante* where to produce. The second is through relocations to another city. Both create challenges for identifying a sorting effect. For instance, it is virtually impossible to distinguish entrants who endogenously choose a particular location from those who exogenously chose a location (e.g., because it was the residence of the business' entrepreneur). Relocations provide a way to control for sorting through the within-establishment changes in density and earnings of relocating establishments. We find that relocations, however, are an order of magnitude less frequent than the entry of new establishments. They also potentially reflect endogenous choices made based on unobservable characteristics.

Nevertheless, we examine relocations and entrants across metropolitan areas to gauge their patterns and potential effect on the estimated density premium. Within-establishment estimates of the density premium that control for sorting through relocation imply that up to 90 percent of our original estimated density premium. Because of the aforementioned measurement issues, we consider this to be an upper bound. Looking at actual relocation patterns across metropolitan areas, we find that the majority of relocations are done by establishments with the highest average earnings. Furthermore, we find that these establishments disproportionately relocate towards *lower-density* metropolitan areas. This suggests sorting way work to understate the density premium, though the aggregate effect is likely small because of the infrequency of relocations. It also suggests that the "nursery city" model of Duranton and Puga (2001) may best describe establishment relocation patterns across cities. In their model, firms first locate in more diverse rather than denser cities to find their optimal production process through experimentation. Like dense cities, diverse cities in their model have higher congestion costs. Consequently, once firms learn their ideal process, they relocate to more specialized (and plausible less dense) cities to avoid higher congestion costs.

Examining the differences in the average earnings of entrants across metropolitan areas, we find that entrants in high-density cities have earnings that are 21 percent higher, on average, than entrants in low-density cities. Relative to the differences between incumbent establishments in their respective cities, the difference in average earnings is negligible. We also focus on differences between entrants from multi-unit firms, on the working hypothesis that such firms are more likely to endogenously choose the location of their establishments. Therefore, differences among these entrants should be larger if sorting at entry is important. We find, however, that the difference in average earnings between these groups is smaller, at 16 percent, and that relative to incumbents, entrants in low-density metropolitan areas have higher average earnings.

Overall, the data reject a role for learning, selection, and sorting for the estimated density premium, though given challenges with identification, the latter result is more tenuous. This is in stark contrast to similar research for workers that finds steeper earnings-tenure profiles in larger cities and a sizeable role for worker sorting through migration. Finally, we find strong evidence that, for relocating establishments, sorting works in the opposite direction—towards lower-density cities. The finding lends credence to the notion of large, dense cities as "nursery cities" for new firms.

2. Data and Measurement

We use establishment data from the Longitudinal Business Database (LBD) of the Census Bureau.[3] The data include payroll and employment information for nearly every establishment in the U.S. on an annual basis, in addition to a variety of information on the establishment (e.g., industry, location, whether it is part of a multi-establishment firm). We focus on establishments in 1992 and 1997, though we use data from all available years to best identify measures such as entry, exit, and establishment age. We focus on these two years because they are Economic Census years, meaning that

[3] For additional details about the LBD, see the Appendix and Jarmin and Miranda (2002).

the U.S. Census Bureau conducted an extensive census of all businesses, and these years tend to have the best measures of establishment entry and exit. These years are also the last two Census years available that have consistent measures of industry across years; the U.S. changed its classification system from the older Standard Industrial Classification (SIC) system to the North American Industrial Classification System (NAICS) in 1997. We restrict our analysis to private, non-agricultural establishments, giving us 4.9 million observations in 1992 and 5.3 million observations in 1997.

We use the Consolidated Business Statistical Area (CBSA) definition of metropolitan areas as our city-level unit of analysis, focusing only on the Metropolitan Area locations (i.e., we ignore the smaller locations classified as Micropolitan Areas under the CBSA system). This provides us with 363 CBSAs in our study.[4] Our measure of urban density is 1990 population per square mile, which we calculate for each CBSA by aggregating population and land area data up from the county level. We use the same approach to calculate the share of the 1990 CBSA population with a college degree. We use the college share as a proxy for the average worker skill in a CBSA, but note that it is a crude proxy since it will not capture variations in other skills, both observable and unobservable, that are not related to education.

We measure entry and exit at the annual frequency. This ensures that all exits measured in 1992 occurred during that year and all entrants measured in 1997 occurred during that year (rather than during the intervening five-year period). We define an entry (exit) as the first (last) time an establishment appears with positive employment in the available sample of the LBD, which spans 1975 through 2005. We also measure establishment age using the full LBD sample. An establishment is assigned an initial age of zero years at entry. Since we can only identify age via observing the establishment in the LBD, we topcode age at 16 years (the maximum observed age in 1992) for both years in our sample.

---

[4] These CBSAs roughly correspond to the older definitions of Metropolitan Statistical Areas and Primary Metropolitan Statistical Areas.

We use payroll per employee as our measure of average earnings at each establishment. Doing so confronts us with several measurement issues. First, payroll in the LBD covers all individuals paid during the year but employment is reported for a particular point in time (March of the year). Thus, a standard measure of payroll per employee could tend to overstate average earnings for establishments that had high worker turnover or were rapidly contracting during the year, and tend to understate earnings for establishments who were rapidly expanding during the year. Second, there is the timing of the payroll and employment data. Payroll in the LBD covers all employees paid during the calendar year (January to December). However, employment is measured during the year (in March). Finally, measurement error in either payroll or employment could lead to extreme outliers in the average earnings measure.

To obtain a more accurate measure of earnings, we define the average earnings for an establishment in year t as the total annual payroll in year $t-1$ divided by the average of employment in years $t-1$ and t, or

(1) $$w_{et} = \frac{P_{e,t-1}}{\frac{1}{2}(N_{et}+N_{e,t-1})}$$

where $P_{e,t-1}$ is the total annual payroll of establishment e in year $t-1$ and N_{et} is the reported employment of establishment e in year t. We define the average earnings of entrants as P_{et}/N_{et} and the average earnings of exits as $P_{e,t-1}/N_{e,t-1}$. We then evaluate these measures for outliers and impute an average earnings measure where necessary. We detail our evaluation an imputation algorithm in the appendix. Finally, we deflate our earnings measures using the Consumer Price Index to 1997 dollars.

Throughout our analysis, we consider the average earnings of an establishment as a proxy for its productivity. There are several issues with this. Earnings are a cost to the firm as much as they are a rent paid to productive labor. In addition, average earnings represent an average across a distribution of workers while productivity is usually thought of as a uniform measure within an establishment.

The LBD does not allow a direct measure of productivity, but we check and reaffirm the validity of earnings as a proxy in several ways. First, we note that our results below on the density premium across the earnings distribution parallel the findings of Combes et al. (2008), who perform a similar exercise with TFP. Namely, we find an urban density premium that increases with average earnings while Combes et al. find a city size premium that increases with TFP. Second, our results below on establishment exit show that establishments exit rates declines sharply with average establishment earnings. This runs counter the concern that high-earnings establishments are predominantly high-cost rather than high-productivity establishments.

Finally, we replicate the results of Syverson (2004) using our average earnings measure in lieu of total factor productivity (TFP). The Syverson study is a particularly useful because it focuses on the relationship between establishment-level dynamics and urban density. Syverson suggests that locations with a greater density of demand will have greater competition among local firms, and consequently, greater exit. He tests and affirms the main implications of his model by looking at differences in the TFP distributions of plants in the ready-mix concrete industry across areas with different construction employment densities (the construction industry is the primary consumer of ready-mix concrete). He focuses on six moments: the (weighted) mean, median, and interquartile range of the TFP distribution; the TFP of the plant at the tenth percentile of the TFP distribution; (log) average plant size; and the producer-demand ratio (the number of plants per 1,000 construction workers). He regresses each moment separately on (log) density. He finds that areas with greater demand density have a less disperse TFP distribution that exhibits greater lower truncation. These areas also have higher average TFP, larger plants, and a lower producer demand ratio.

We replicate the Syverson study using the LBD data an identical subsample: plants in the ready-mix concrete industry (SIC 3273) with at least 5 employees in locations with at least 5 plants for the years 1982, 1987, and 1992. Our analyses differ in only three regards: i) we use data from the LBD rather than from the Census of Manufactures, ii) we use the CBSA definitions rather than the Component Economic Area (CEA) definitions of a metropolitan area, and iii) we use the log of average earnings instead of the log of TFP. The first difference is negligible as, during economic census years, the coverage of the LBD and Census of Manufactures is nearly identical. The main implication of the second difference is that unlike Syverson's analysis, our analysis excludes rural locations (CEA definitions cover non-metropolitan areas). The third difference is the margin of interest. Our results, along with the original results from Syverson (2004), are in Table 1. As one can see, we find qualitatively similar results for all six moments. Thus, we conclude that there is in fact a strong correlation between our measure of average earnings and establishment productivity.

As a last point, it is worth reporting how basic establishment characteristics and behavior vary with urban density. Table 2 presents basic sample statistics as well as the slope coefficients from OLS regressions at the establishment level of the average (log) number of employees and age of establishments, the average annual entry rate, and the average annual exit rate on (log) density using observations from the 363 CBSAs in our sample. We report the results from the regression on density alone, a regression that controls for the share of the CBSA population that is college educated, and a regression that additionally controls for establishment characteristics (size, age, industry, and multi-unit firm status, excluding the characteristic used as the dependent variable).[5] The summary statistics show considerable dispersion in establishment size, age entry and exit rates across establishments, and to a lesser degree, across metropolitan areas. The regression results show that more dense CBSAs tend to have smaller but older establishments, on average. Establishment entry rates decline with density, while exit rates are essentially unrelated with density, especially when controls for the CBSA college share and the remaining establishment characteristics are added.

[5] All regressions include a dummy variable for year. In these and all subsequent establishment-level regressions in the paper, we report standard errors that are clustered at the CBSA level.

3. Basic Evidence on the Establishment Characteristics and Density

As Figure 1 showed earlier, we find an elasticity of average earnings with respect to density at the CBSA level of 8.1 percent, with a standard error of 0.7 percent. Controlling for college share only reduces the elasticity to 7.8 percent (with a standard error of 0.7 percent).[6] These estimates are roughly in line with estimates previously found using data on individuals rather than establishments.[7]

Figure 1 reports the CBSA-level elasticity of wages with respect to density. A richer estimation at the establishment level allows us to control for various establishment characteristics, such as size, age, detailed industry, and whether the establishment is part of a multi-unit firm. It is well known that earnings vary strongly with industry, size, and age (for example, see Brown and Medoff, 1989, 2003). We regress log average earnings for each establishment on our density measure, with and without the above controls. We use both years of our panel and cluster our standard errors by CBSA. We also examine the micro-level relation controlling for the CBSA college share. Specifically, for establishment e in CBSA j at year t, we estimate

(2) $$\ln w_{ejt} = \alpha_t + \beta \ln D_j + \gamma C_j + \delta Z_{et} + \varepsilon_{ejt},$$

where $\ln w_{ejt}$ is the log of average establishment earnings, $\ln D_j$ is our density measure, C_j is the CBSA college share, Z_{et} is the set of establishment controls (the log of employment, age, four-digit SIC industry, and an indicator for membership in a multi-unit firm), and α_t is a year dummy.

The results for the full sample of establishment-years appear in Table 3. Unconditionally, we find a somewhat higher density premium at the establishment level relative to using aggregate data, 10.2 percent. Controlling for the CBSA college share reduces this estimate to 8.0 percent. Controlling for both college share and establishment characteristics reduces the estimate further, to 7.4 percent.[8] In each

---

[6] The coefficient on the college share is 0.33 with a standard error of 0.10.
[7] See, for example, Ciccone and Hall (1996) and Sveikauskas (1975).
[8] This is consistent with recent work by Lehmer and Möeller (2010), who find an urban density premium that persists after controlling for firm size.

case, the estimated coefficients are highly significant. Thus, even at the micro level, and even after controlling for establishment characteristics that are known to exhibit strong correlations with earnings, we still find a large and significant density premium for establishments.

It is plausible that the above estimates mask wide heterogeneity in the density premium across different subgroups of the data. In Table 4, we re-estimate (2) separately for different subgroups of the data. The first groups we examine are entering and exiting establishments.[9] Entrants and exits each exhibit a slightly higher elasticity of earnings with respect to density relative to all establishments, but neither of their coefficients on density is significantly different from their counterpart in Table 3. We next estimate the density premium separately for multi-unit and single-unit firms. Of all the different cuts of the data, single-unit versus multi-unit firms is the only grouping where we find a significant difference in the estimated elasticity. Single-unit firms exhibit the higher density premium, 8.0 percent versus 5.8 percent for multi-unit firms.[10] We also estimate the density premium for five establishment size classes and for five broad industry groups (construction, manufacturing, retail trade, finance and professional services, and local services). Across size classes, we find no significant differences in the earnings premium across groups. Across industries, finance and professional services has the highest estimated density premium, while retail trade and local services have the lowest estimates.

Estimates of an urban density premium face an endogeneity issue: urban density may be a consequence rather than a cause of local productivity advantages. To deal with this, we re-estimate our (2) using an two-stage least squares approach where we instrument the density and college share variables with a variety of geological data for each CBSA.[11] This mimics the approach taken by Combes et

[9] We exclude age as a control for establishment characteristics in the entry regressions since, by definition, all entrants are zero years old.

[10] It is worth noting that this finding is consistent with research by Henderson (2003), who found that measures of localization and urbanization economies generated higher estimated returns for single-plant manufacturing firms.

[11] We also replicate our subsequent analyses using the IV approach, and report these results in the appendix. Estimates from the first stage regressions are also in the appendix.

al. (2010). The geology variables include data on the fraction of the CBSA water-covered, the mean elevation and the fraction of the CBSA above 1000m, an index of terrain ruggedness, the average annual temperature and moisture, the number of potential growing days, and the fraction of the soil represented by a vector of soil types. The results are in Table 5. We only have such data for 283 of our 363 CBSAs, so we report both the OLS and IV results for this subsample. The results show that endogeneity does not account for the observed density premium. When one includes the college share and establishment characteristics, the OLS and IV estimates of the density premium are nearly identical.

Finally, we examine whether there are differential returns to urban agglomeration across the earnings distribution. Research on workers' earnings has found that the return to agglomeration is higher for more skilled workers (e.g., Glaeser and Mare, 2001; Gould, 2007; Baum-Snow and Pavan, 2010b; and Glaeser and Resseger, 2010). Combes et al. (2009) have also found that firms with higher TFP have a higher return to agglomeration. One can interpret the findings, in the case of workers, as evidence of "skill-biased" returns to agglomeration, or in the case of firms, as "productivity-based" returns to agglomeration. That is, workers with higher observable skills, or firms that are more productive, tend to reap a greater benefit from locating in a dense city.

Figure 2 shows the full distribution of earnings for two subsets of the data: establishments in the top quartile and in the bottom quartile of CBSAs, ranked by their population density.[12] The data are pooled over both years and the kernel density estimates are based on an unconditional earnings measure. The figure clearly shows a rightward shift of the entire earnings distribution for establishments in the high-density CBSAs. Among continuing establishments, median earnings are 23 percent higher in high-density CBSAs unconditionally, and 14 percent higher when controlling for establishment

[12] The least-dense CBSA in the top quartile is Louisville-Jefferson County, KY-IN and the densest CBSA in the top quartile is Niles-Benton Harbor, MI.

characteristics. The earnings distribution in the densest CBSAs also exhibits greater dispersion. The 90-10 ratio is 9.8 log points larger in high-density CBSAs unconditionally and 14.6 log points larger after controls. With such differences in both the levels and dispersion of average earnings across CBSAs, it is natural to examine whether the establishments experience a greater density premium at different points of the earnings distribution.

We split the earnings distribution of each CBSA into quintiles and estimate the earnings premium separately for establishments within each quintile. We first identify which quintile of their CBSA-specific earnings distribution each establishment belongs to, then group all establishments into five categories based on their CBSA-specific quintile. We then estimate (2) for each group.

Table 6 presents the results for three specifications: a regression on the log of density and controls for year only; a regression that adds the college share as an additional control; and a regression representing the full specification in (2).[13] Across all specifications, the estimates show a clear monotonic rise in the estimated density premium with average establishment earnings across the distribution. Without additional controls, the difference in the estimated elasticity between the lowest and highest quintile is 6.4 percentage points, with the estimate in the highest quintile (14.4 percent) about 80 percent higher than that of the lowest quintile (8.0 percent). With all controls included, the spread is 3.5 percentage points, and the estimated premium in the highest quintile (10.2 percent) is about 52 percent higher than the estimated premium in the lowest quintile (6.7 percent). These results are broadly consistent with the findings of Combes et al. (2009). They find that the returns to TFP of locating in a large city are 14 percent for the top quartile of firms and 5 percent for the bottom quartile. Thus, like Combes et al., we find strong evidence of "productivity-biased" returns to agglomeration. We

[13] We also experimented with an alternative estimation strategy that regressed log earnings on the control variables and then estimated the density premium after re-sorting establishments across the quintiles based on the within-CBSA ranking of their resulting *residual* earnings measure. The approach produced very similar results to the ones reported in Table 6.

do so despite using average earnings rather than TFP as a productivity measure and despite examining urban density rather than city size.

4. The Roles of Learning, Selection, and Sorting

We next test for evidence of learning, selection, and sorting by establishments. Previous research on workers have found that their city tenure-wage profiles tend to rise more steeply in denser cities (e.g., Glaeser and Mare, 2001; Baum-Snow and Pavan, 2010a; de la Roca and Puga, 2011). One can interpret these findings as evidence of faster human capital accumulation, or a faster degree of "learning." We examine whether there is also evidence of "learning" (e.g., through knowledge spillovers or learning about customer demand) for establishments. Firm selection (through exit) and sorting (through endogenous location choice) can lead to an over-estimate of the density premium for establishments. Firm sorting refers to the self-selection of high-productivity firms into dense locations. Firm selection refers to the lower-truncation of the firm productivity distribution through the exit of low-productivity firms (i.e., negative selection). Such a selection effect can be stronger in denser areas because of greater competition. The stronger selection effect will lead to a greater lower-truncation of the productivity distribution, and a tighter correlation between density and average firm productivity that is unrelated to any firm returns to agglomeration.

4.A. Learning: The Density Premium and Age

We next explore whether the density premium rises with an establishment's tenure at a particular location, which we would attribute to faster learning in denser cities. Unlike workers, firms rarely migrate across areas. As we show below, the propensity of establishments to relocate across cities is about one-tenth as common as the entry of a newly-formed establishment into a city. Thus, for nearly all establishments, their city tenure equals their age. Therefore, we examine whether the average establishment earnings increases with age faster in denser cities. If so, the evidence would suggest that firms, like workers, accumulate returns to agglomeration over time. This may occur through greater flexibility in experimenting with production processes, through knowledge spillovers accumulated over time, or through learning about the local demand over time.

To test for evidence of learning, we repeat our establishment-level regressions of (log) earnings on population density including fixed effects for the age of the establishment and an interaction between these fixed effects and density. As before, we also run additional specifications that control for the CBSA college share and establishment characteristics. The general form of the regression we run is

(3) $$\ln w_{ejt}(a) = \alpha_t^1 + \varphi(a) + \beta^1 \ln D_j + \zeta(a) \ln D_j + \gamma^1 C_j + \delta^1 \tilde{Z}_{et} + \varepsilon_{ejt}^1(a),$$

where a denotes age and $\tilde{Z}_{et}$ represents the same establishment characteristics as before except for age (industry, size, and multi-unit firm status). Standard errors are clustered by CBSA.

Figure 3 plots the predicted value of earnings from the baseline specification of (3) that does not control for college share or the establishment characteristics. Specifically, it plots $\hat{\varphi}(a) + \left(\hat{\beta}^1 + \hat{\zeta}(a)\right) \ln D_j$, with density evaluated at its value for the CBSAs ranked at the 90th, 50th, and 10th percentiles of the population density distribution.[14] The figure shows that earnings are higher in more dense areas and that earnings rise with establishment age. Notably, there is no evidence of fanning out of the earnings-age profiles.

The coefficient of interest in this exercise is $\zeta(a)$, since $\partial\hat{\zeta}(a)/\partial a > 0$ would imply a density premium that rises with establishment age, and be suggestive of returns to agglomeration for establishments that accumulate over time. Figure 4 shows that, in fact, the density premium for establishments is essentially flat over their lifespan. It plots the coefficients $\hat{\beta}^1 + \hat{\zeta}(a)$ as a function of

[14] These CBSAs correspond to the Santa Cruz-Watsonville, CA CBSA, the Des Moines-West Des Moines, IA CBSA , and the Yakima, WA CBSA, respectively. While not reported, the predicted earnings estimates for the two specifications with additional controls produce qualitatively similar results.

age for three specifications derived from (3). Across all specifications, the estimated elasticity with respect to density is essentially independent of age. In the baseline specification, the estimated elasticity varies within a relatively tight range of 9.3 to 11.1 percent. When we include all controls, the range is even tighter, between 6.9 and 7.8 percent. In comparison, both of these ranges are smaller than the marginal effect of controlling for the CBSA college share, which reduces the elasticity estimate across all ages by 2.5 percentage points. Further, these differences in the estimated density premium are all well within the standard error bands for their respective specification.

Figure 5 replicates the exercise for different subgroups of the data. The five panels of the figure report the coefficients $\hat{\beta}^1 + \hat{\zeta}(a)$ from the estimation of the full specification in (3) separately by continuing versus exiting establishments, by establishments in multi-unit and single-unit firms, by establishment size class, by major industry group, and by within-CBSA earnings quintile. The results show that both surviving and exiting establishments exhibit a similar density premium regardless of age. Multi-unit firms appear to exhibit a declining density premium with age, but the decline is both statistically insignificant and economically small; young establishments of multi-unit firms exhibit a premium between 6 and 7 percent while the oldest establishments exhibit a premium of about 5 percent. The density premium shows no clear relation to age across our five establishment size classes. Furthermore, the estimates with respect to age are very imprecisely measured for the two largest size classes. There appears to be little change in the estimated density premium in our four of our five industries. Construction is the exception, which shows a premium that rises from 7.1 percent to 11.1 percent as establishments age. Finally, there appears to be no relationship between age and the density premium within earnings quintiles of the earnings distribution within CBSAs. Establishments in the highest quintile show a slight decline among older establishments and establishments in the lowest quintile show a slight decline among younger establishments, but again, neither decline is either statistically or economically significant.

We should note that the estimated density premium from (3) is a combination of a pure establishment-specific return to agglomeration and an average return to agglomeration of its workers. The average return to workers, in turn, is dependent on the turnover rate of the workforce. In the extreme case of an establishment that retains the same workforce throughout its life, the average return to workers should rise with establishment age because, based on earlier research, the density premium earned by workers rises with their city tenure. In the absence of worker turnover, our findings in Figure 4 would suggest that the establishment-specific density premium is *declining* with age. Worker turnover complicates the interpretation. We feel that a plausible assumption is that higher-tenure workers are replaced by lower-tenure workers, on average. If this turnover process were to keep average tenure roughly constant over an establishment's life, then our results would imply a constant establishment-specific return. For there to be an increasing establishment-specific return over time, one would need average worker tenure to be decreasing over the establishment's life. Estimating the path of average worker tenure within the establishment requires access to matched employer-employee data, and we plan to explore this further with such data in future work.

4.B. Selection through Exit

Recent research has tried to quantify how much of the estimated returns to agglomeration in previous studies are due to selection through firm exit. Combes et al. (2009) find a strong selection effect for establishments in French data, but find that it does not vary with city size. They conclude that selection does not account for their estimated returns to agglomeration as a result. Holmes, Hsu, and Lee (2010) make the point, however, that their approach does not give an explicit role to exit, which is a crucial part of standard models of firm selection (see Jovanovic, 1982; Hopenhayn, 1992; and Ericson and Pakes, 1995).

Consequently, we proceed with an analysis of the role of selection by explicitly studying establishment exit patterns as a function of their average earnings and CBSA density. We group all establishments into one-percentile bins based on their ranking within their CBSA-specific earnings distribution. We then pool all establishments based on whether their establishment resides in a CBSA within the top or bottom quartile of the CBSA density distribution. Finally, we calculate the fraction of establishments who exit within each percentile of the earnings distribution for the CBSAs in the top and bottom quartiles of the density distribution. If selection through exit is an important estimating the urban density premium for establishments, we would expect to see significantly higher exit rates for within high-density CBSAs, and see a greater difference towards the left tail of the earnings distribution.

We report results that use an exit probability and average earnings measure that control for establishment characteristics (industry, size, age, multi-unit firm status). This eliminates differences in exit rates due to observable differences in establishment composition across cities. Our estimates are in Figure 7.[15] The left panels show the results for all establishments and the right panels show the results for establishments aged 5 years or less. The latter results are of interest because selection effects may be most important early in an establishment's life cycle, when exit rates are highest. The top panels report exit rates a function of the earnings distribution, while the bottom panels report the difference between exit rates in high-density versus low-density CBSAs (with 95 percent confidence intervals).

The results are consistent with the conclusions of Combes et al. (2009). There is clearly a strong selection effect for establishments. Exit rates decline with average establishment earnings, and the highest exit rates are for establishments in the bottom 20 percent of the earnings distribution. At the same time, the results suggest that there is little role for selection in accounting for the estimated density premium in the data. There is little difference in exit rates between high-density and low-density

[15] We also performed the exercise using the raw earnings and exit probabilities and obtain results that are very similar to those in Figure 7.

CBSAs. For all establishments, exit rates are somewhat higher in high-density CBSAs, but this difference is only marginally significant for the middle of the earnings distribution (between the 30th and 55th percentiles). For younger establishments, the differences are noisier but smaller, on average, and not significant anywhere along the earnings distribution.

[TO BE COMPLETED.]

4.C. Sorting through Relocation

The estimated density premium for establishments can also be affected by sorting. That is, the premium may be overstated because productive establishments sort into dense areas. Studies that focus on the density premium for workers can control for such sorting by estimating the within-worker density premium for individuals who migrate across cities. Examining sorting for establishments is more complex because establishments can sort along two margins. First, existing establishments can relocate from one city to another. So long as one has longitudinal information on the establishment location, one can apply the same approach to these establishments that is used for migrating workers. Issues arise, however, because relocations are relatively rare (1 percent of establishments per year relocate), and because relocating establishments may be *ex ante* different from establishments that do not move. Second, establishments can choose which city to locate in at entry. About 10 percent of all establishments are new entrants in an average year, but identifying entrants who endogenous choose a location from entrants that randomly locate in a particular area (e.g., because it is the home location of the entrepreneur) is virtually impossible.

Because of these limitations, little empirical research has been done on the sorting of firms into cities (an exception is Duranton and Puga, 2001, using French data). In the remainder of this study, we present evidence on the characteristics, behavior, and estimated density premium of relocating establishments and the characteristics of entrants across CBSAs.

We start with relocating establishments. We identify relocations in the data as establishments who change their county code no more than once during their existence. We ignore multiple moves because of the potential for measurement error in location codes (e.g., coding errors that are corrected in a subsequent year). We also focus only on continuing establishments over the 1991-92 and 1996-97 periods, since an establishment must be in existence for both years to move. Finally, we only examine moves between CBSAs, since the measure of interest is the change in metropolitan area population density. This identifies just over 81,000 moves, which represents 1.0 percent of all continuous establishments.[16]

Table 8 reports the differences in the basic characteristics of relocating versus static establishments. Relocating establishments tend to be younger and slightly larger, on average. They also tend to move to less dense CBSAs that have somewhat lower shares of their population with a college degree. [TO BE COMPLETED.]

Table 9 reports the results of the earnings-density regression in (2), with and without establishment and college share controls, in levels and in first differences. The levels specification replicates the results of Table 3 for all continuous establishments. The first difference specification identifies a density premium through changes in CBSA density among relocating establishments. Estimates from the first-difference specification suggest that about 90 percent of the estimated density premium is due to sorting. Without controls, the estimate falls from 10.1 percent to 0.8 percent. With controls, the estimate falls from 7.2 percent to 0.7 percent. We consider this an upper bound on the role of sorting, however, because establishments who relocate may be precisely those who do not gain much from locating in a dense location.

[16] Inter-CBSA moves represent only 35 percent of all inter-county moves, but 79 percent of all moves out of a metropolitan area (the remainder are moves to non-metropolitan areas). We reject about 26 percent of potential relocations because of multiple changes to an establishment's county code.

Because of this issue, we perform a more in-depth examination of relocation patterns as a function of average establishment earnings and CBSA density. As we did with exit rates, we order establishments based on their percentile within their CBSA earnings distribution. We do this based on their "origin CBSA" (their location in 1991 or 1996) and their "destination CBSA" (their location in 1992 or 1997). For each percentile of the origin (destination) CBSA distribution, we calculate the fraction of establishments that moved out of (into) the CBSA. As before, we pool CBSAs based on their quartile of the CBSA density distribution and report relocation rates for the highest and lowest quartile. We also examine estimates based on raw earnings and relocation measures and measures that control for establishment characteristics, though we report only the latter.[17]

The results are in Figure 8. The top panels show the relocation rates into and out of low-density CBSAs (left) and high-density CBSAs (left). The bottom panels show the difference between high-density and low-density out-migration (left) and in-migration. Several things stand out in the figure. First, high-earnings establishments are the most likely to relocate. Second, across the earnings distribution, there is net out-migration in high-density CBSAs and net in-migration in low density CBSAs. Third, the net migration toward low-density CBSAs is most pronounced and statistically significant for the highest-earnings establishments (those in the top 30 percent of their destination-CBSA earnings distribution). Fourth, out-migration from low-density CBSAs is essentially uncorrelated with average establishment earnings. This is not true for out-migration from high-density CBSAs, which rises sharply with earnings at the right tail of the distribution.

Taken together, the results suggest that sorting based on relocation works to *reduce* the estimated density premium, since the most productive establishments are likely to relocate to *less* dense CBSAs. The total effect of this type of sorting is likely small, however, since it involves only 1

[17] As before, results using the raw measures produce similar results to those using the conditional measures.

percent of continuing establishments. It also suggests that the subset of movers may indeed be different in from stayers. Putting together the evidence from Figure 8 and Table 8 suggests that movers tend to be young and highly productive, and that they enter dense locations then systematically move to less dense locations after some time.

This pattern is consistent with at least one theory of agglomeration and firm dynamics. Specifically, Duranton and Puga (2001) propose a model where firms prefer to initially enter more diverse cities. These cities are more congested, and therefore more costly to operate in, but firms do not initially know their ideal production process, and the diversity these cities afford allow for experimentation. Over time, firms learn their ideal process, and the benefit of diversity is outweighed by the congestion costs. Consequently, firms that successfully implement their ideal process eventually move from the diverse "nursery cities" to specialized cities where congestion costs are lower.

While our evidence pertains to dense rather than diverse cities, the general pattern we observe is consist with the Duranton-Puga model, partly because dense cities also tend to be larger and industrially diverse. When considering our complete, results, however, tensions exist with the model and the evidence. Namely, we also find evidence of "productivity-biased" returns to density. If the most productive establishments reap the greatest density premium, then why is it the case that they are also the most likely to move to a lower-density city? We plan to address this tension in future research.

4.D. Sorting through Entry

As a final exercise, we document differences in the earnings distribution of entrants between high-density and low-density CBSAs. As we noted earlier, identification of sorting through entry from the random allocation of entrants into dense cities (perhaps based on the residence of the entrepreneur) is virtually impossible. Consequently, we focus on examining earnings differences among entrants to see if there is at least first-order evidence of sorting through entry.

We examine differences in the distribution of average establishment earnings across high-density and low-density metropolitan areas for entrants in absolute terms and relative to incumbent establishments. We also examine differences among the new establishments of multi-unit firms. The hypothesis is that multi-unit firms are much more likely to endogenously choose the location of a new establishment from a set of metropolitan areas, whereas a new, single-establishment firm is more likely to start up where its entrepreneur is currently located. Therefore, if there is sorting into more dense locations by more productive establishments, it should be more pronounced for the new establishments of multi-unit firms.

Figure 9 plots kernel density estimates of the earnings distributions of all entrants separately for those located in CBSAs ranked in the top and bottom quartile of the population density rankings. The estimates do not control for differences in establishment characteristics. The figure shows that the earnings distribution of entrants in the high-density CBSAs has higher mean earnings and exhibits greater dispersion. Table 10 presents the distributional statistics for entrants in each CBSA group after conditioning establishment characteristics out of the earnings measure. Entrants in high-density CBSAs have earnings that are 21.2 percent higher, on average, than the earnings of entrants in low-density CBSAs. The difference in median earnings is nearly as large (18.0 percent).[18] The earnings distribution of entrants also has greater dispersion in high-density CBSAs. An issue with the raw comparison of earnings distributions is that there is a similar disparity between the earnings distribution of incumbent establishments in high-density and low-density CBSAs. We therefore report the difference-in-difference estimate of the entrant earnings distributions relative to the incumbent earnings distributions in the last column of Table 10. When we do so, the relative difference between entrant earnings distributions is only 1.3 percent, on average, and essentially zero at the median.

[18] Without controlling for establishment characteristics, the difference in mean earnings is 26.0 percent and the difference in median earnings is 24.2 percent.

If our working hypothesis that entrants of multi-unit firms are more likely to sort is correct, then the evidence on the earnings differences among this subset of entrants does not suggest a strong role for sorting. Mean earnings for multi-unit entrants are 16.3 percent higher, and median earnings are 13.9 percent higher, in high-density CBSAs. Furthermore, relative to incumbents in their own CBSA, entrants of multi-unit firms have mean and median earnings that are 4.9 and 4.1 percentage points lower, respectively, in high-density CBSAs. These nonparametric results are not a clear rejection of a role for sorting by any means. At the same time, they do not present strong first-order evidence for the sorting of high-earnings entrants into high-density CBSAs.

5. Conclusions

This paper presents new evidence on the earnings premium associated with urban density for establishments. It generates new stylized facts on the relation between average establishment earnings, which we consider a proxy for productivity, and urban density that are analogous similar facts other studies have found for the earnings of workers. We find a density premium of about 7.4 percent after controlling for the college share of a metropolitan area and observable establishment characteristics. The estimate varies little across various subgroups of the data and is robust to controlling for the joint endogenous determination of density and productivity. Analogous to Combes et al. (2009), we also find that the density premium rises with average establishment earnings.

This paper also presents evidence related to three facets of urban agglomeration that has been relevant for both workers and firms in the literature: greater knowledge accumulation in dense cities, greater selection through exit, and greater sorting of productive agents into dense cities. To varying degrees, we find evidence against all three. We show that, for establishments, the urban density premium does not rise with establishment age (or equivalently, city tenure), implying that any potential returns to agglomeration are realized at entry. We also find no variation in a selection effect across metropolitan areas. Exit rates follow a similar pattern across the earnings distribution for both high-density and low-density metropolitan areas. Finally, we find a strong role for based on estimates identified through establishment relocations, but at the same time, we find that relocating establishments tend to move to *lower-density* metropolitan areas. Entrants in high-density cities have higher earnings than their counterparts in low-density cities in absolute terms, but not relative to incumbent establishments their respective cities.

Our results shed new light on how establishment dynamics relate to the urban density and provide new guidance for theories of urban agglomeration. Most notably, we find evidence of a "nursery city" effect for dense areas. The young and most productive establishments in a high-density area are most likely to relocate, and they tend to move to a lower-density area when they do so. This behavior is consistent with the model of Duranton and Puga (2001), though they focus on urban diversity rather than density. Future work based on a nursery city model still faces challenges from our other evidence, however. Notably, it is unclear why the most productive establishments move to lower-density cities when we, along with Combes et al. (2009) find evidence that the returns to agglomeration are highest for these establishments. There is also further research needed on the role of firm learning. Under the assumption that average worker tenure is nondecreasing within the establishment over time, our estimates of the density premium are an upper bound on the establishment-specific return to density with respect to age. A more detailed analysis would require matched-employee micro data that would allow one to account for the tenure and mobility of workers within the establishment. A finding of an establishment-specific effect that steadily decreases with age, for example, may help reconcile the relocation behavior of high-earnings establishment with their high estimated returns to density. These establishments may eventually relocate because their returns slowly erode over time.

References

Abowd, John M., Patrick Corbel, and Francis Kramarz, 1999. "The Entry and Exit of Workers and the Growth of Employment: An Analysis of French Establishments," *Review of Economics and Statistics,* 81(2); 170-87.

Bacolod, Marigee, Benardo Blum, and William Strange, 2009. "Skills and the City," *Journal of Urban Economics,* 65(2), pp. 127-135.

Baum-Snow, Nathaniel, and Ronni Pavan, 2010a. "Understanding the City Size Wage Gap," Brown University, mimeo.

Baum-Snow, Nathaniel, and Ronni Pavan, 2010b. "Inequality and City Size," Brown University, mimeo.

Behrens, Kristian, Gilles Duranton, and Frederic Robert-Nicoud, 2010. "Productive Cities: Sorting, Selection, and Agglomeration," CEPR Discussion Paper No. 7922.

Brown, Charles, and Medoff, 1999. "The Employer-Size Wage Effect," *Journal of Political Economy,* 97(5): 1027-59.

Brown, Charles, and James L. Medoff, 2003. "Firm Age and Wages," *Journal of Labor Economics*, 21(3): 677-97.

Ciccone, Antonio, and Robert E. Hall, 1996. "Productivity and the Density of Economic Activity," *American Economic Review,* 86(1): pp. 54-70.

Combes, Pierre-Philippe, Gilles Duranton, Laurent Gobillon, Diego Puga, and Sébastien Roux, 2009. "The Productivity Advantages of Large Cities: Distinguishing Agglomeration from Firm Selection," CEPR Discussion Paper No. 7191.

Combes, Pierre-Philippe, Gilles Duranton, Laurent Gobillon, and Sébastien Roux, 2010. "Estimating Agglomeration Economies with History, Geology, and Worker Effects," in *Aggomeration Economics*, Edward L. Glaeser (ed.), University of Chicago Press: Chicago, IL.

de La Roca, Jorge, and Diego Puga, 2010. "Learning by Working in Dense Cities," IMDEA, mimeo.

Duranton, Gilles, and Diego Puga, 2001. "Nursery Cities: Urban Diversity, Process Innovation, and the Life Cycle of Products," *American Economic Review*, 91(5): 1454-77.

Ericson, Richard, and Ariel Pakes, 1995. "Markov-Perfect Industry Dynamics: A Framework for Empirical Work," *Review of Economic Studies,* 62(1): 53-82.

Freedman, Matthew, 2008. "Job Hopping, Earnings Dynamics, and Industrial Agglomeration in the Software Publishing Industry," *Journal of Urban Economics*, 64(3), pp. 590-600.

Glaeser, Edward L., 1999. "Learning in Cities," *Journal of Urban Economics,* 46(2): pp. 254-77.

Glaeser, Edward L., and David C. Mare, 2001, “Cities and Skills,” *Journal of Labor Economics*, 19(2): pp. 316-42.

Glaeser, Edward L., and Matthew G. Resseger, 2010. “The Complementarity between Cities and Skills,” *Journal of Regional Science,* 50(1): 221-44.

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

Henderson, J. Vernon, 2003. “Marshall’s Scale Economies,” *Journal of Urban Economics*, 53(1): 1-28.

Holmes, Thomas, Wen-Tai Hsu, and Sanghoon Lee, 2010. “A Model of Cities, Entrepreneurship, and Exit,”University of Minnesota, mimeo.

Hopenhayn, Hugo, 1992. “Entry, Exit, and Firm Dynamics in Long Run Equilibrium,” *Econometrica,* 60(5): pp. 1127-50.

Jarmin, Ron, and Javier Miranda,2002. “The Longitudinal Business Database,” Center for Economic Studies Working Paper no. 02-17, U.S. Census Bureau.

Jovanovic, Boyan, 1982. “Selection and the Evolution of Industry,” *Econometrica*, 50(3), 649-70.

Lehmer, Florian, and Joachim Möeller, 2010. “Interrelations between the Urban Wage Premium and Firm-Size Wage Differentials: A Microdata Cohort Analysis for Germany,” *Annals of Regional Science*, 45(1), pp. 31-53.

Moretti, Enrico, 2004. “Workers’ Education, Spillovers, and Productivity: Evidence from Plant-Level Production Functions,” *American Economic Review*, 94(3): pp.656-90.

Spletzer, James, R. Jason Faberman, Akbar Sadeghi, Richard Clayton, and David Talan, 2004. “Business Employment Dynamics,” *Monthly Labor Review*, 127(4): pp 29-42.

Sveikauskas, Leo, 1975. “The Productivity of Cities,” *Quarterly Journal of Economics,* 89(3): pp. 319-413.

Syverson, Chad, 2004. “Market Structure and Productivity: A Concrete Example,” *Journal of Political Economy* 112(6): 1181-1222.

Wheeler, Christopher, 2001. “Search, Sorting, and Urban Agglomeration,” *Journal of Labor Economics*, 19(4): pp. 879-99.

Table 1. Comparison of Results for Firm Selection in the Concrete Industry

Moment	*Estimate of Demand Density Elasticity* Estimate from Syverson (2004), using TFP for y_{et}	Estimate from the LBD, using avg. earnings for y_{et}
Interquartile range of distribution of $\ln y_{et}$	-0.015 (0.004)	-0.028 (0.013)
Median value of $\ln y_{et}$	0.018 (0.003)	0.095 (0.015)
Size-weighted mean of $\ln y_{et}$[1]	0.024 (0.004)	0.081 (0.015)
Tenth percentile of distribution of $\ln y_{et}$	0.056 (0.010)	0.080 (0.027)
Mean plant size[1]	0.211 (0.012)	0.065 (0.016)
Producer-demand ratio[2]	-0.363 (0.015)	-0.680 (0.033)
Number of Observations	665	410

Notes: Table reports the estimates from the regression of the listed moment on a measure of demand density (the log of construction employment) and a year dummy across geographic locations (BEA Census Economic Areas for Syverson, and our sample of CBSAs for the LBD.) Estimates in the first column come from Syverson (2004, "Model 2" on p. 1206.), and estimates in the second column are authors' estimates from the LBD. See text for more details. Standard errors are in parentheses.

1. Size is measured as the log of total sales in Syverson (2004) and as the log of employment in the LBD.
2. The producer-demand ratio is the number of plants per 1,000 construction employees.

Table 2. Basic Statistics on Relationships between CBSA Establishment Characteristics and Density

	ln *Size* (employees)	Age (years)	Entry Rate (share of estabs.)	Exit Rate (share of estabs.)
Sample Mean	1.50	8.01	0.103	0.092
Std. Deviation across Establishments	1.37	6.80	0.304	0.289
Std. Deviation across CBSAs	0.10	0.86	0.009	0.015
OLS regression on ln (Density)				
$\ln D_j$	-0.023	0.101	-0.003	0.002
	(0.021)	(0.044)	(0.001)	(0.001)
R^2	0.000	0.001	0.000	0.000
OLS regression on ln(Density) and College Share				
$\ln D_j$	-0.018	0.180	-0.004	0.001
	(0.022)	(0.053)	(0.001)	(0.001)
College Share, C_j	-0.252	-3.735	0.065	0.018
	(0.150)	(0.841)	(0.018)	(0.013)
R^2	0.000	0.002	0.000	0.000
OLS regression on ln(Density) and College Share				
$\ln D_j$	-0.007	0.189	-0.004	0.000
	(0.012)	(0.052)	(0.001)	(0.002)
College Share, C_j	-0.186	-2.139	0.034	0.020
	(0.090)	(0.672)	(0.015)	(0.012)
Establishment Controls	Age, Multi Status, Industry	Size, Age, Multi Status, Industry	Size, Multi Status, Industry	Size, Age, Multi Status, Industry
R^2	0.332	0.195	0.087	0.131
Number of Observations	10,256,604			

Notes: Table reports summary statistics for the listed variables in each column, as well as the results of regressions of the listed variables on the log of 1990 population density and the share of the 1990 population with a college degree. All regression specifications include a year dummy. Establishment characteristics, where listed, include the log of establishment employment, a dummy for whether the establishment is part of a multi-unit firm, fixed effects for age, and fixed effects for four-digit SIC. Standard errors, clustered by CBSA, are in parentheses.

Table 3. Establishment-Level Relations between Earnings and Density

	All Establishments			
	(1)	(2)	(3)	(4)
$\ln D_j$	0.102	0.080	0.092	0.074
	(0.007)	(0.010)	(0.010)	(0.010)
College Share, C_j		1.024		0.883
		(0.102)		(0.093)
Year effects?	Yes	Yes	Yes	Yes
Controls for establishment characteristics?	No	No	Yes	Yes
R^2	0.014	0.017	0.310	0.313
Number of Observations	10,256,604			

Notes: Table reports estimates from the regression of the log of average establishment earnings on the listed variables for our sample of establishment-year observations from the LBD. Establishment characteristics include the log of establishment employment, a dummy for whether the establishment is part of a multi-unit firm, fixed effects for age, and fixed effects for four-digit SIC. Standard errors, clustered by CBSA, are in parentheses.

Table 4. Establishment-Level Relations between Earnings and Density by Sub-Group

	Entrants and Exits		***Multi- & Single-Unit Firms***	
	Entrants	Exits	Single-Unit	Multi-Unit
$\ln D_j$	0.076 (0.011)	0.079 (0.013)	0.080 (0.010)	0.058 (0.009)
College Share, C_j	1.135 (0.129)	0.815 (0.100)	0.945 (0.108)	0.700 (0.088)
Year effects?	Yes	Yes	Yes	Yes
Controls for establishment characteristics?	Yes	Yes	Yes	Yes
R^2	0.257	0.271	0.279	0.460
Number of Observations	1,063,789	950,456	7,587,861	2,668,379

	By Establishment Size				
	1 to 9 Employees	10 to 99 Employees	100 to 249 Employees	250 to 999 Employees	1,000+ Employees
$\ln D_j$	0.079 (0.010)	0.064 (0.010)	0.067 (0.009)	0.075 (0.012)	0.071 (0.013)
College Share, C_j	0.926 (0.104)	0.769 (0.084)	0.862 (0.111)	0.749 (0.113)	0.703 (0.421)
Year effects?	Yes	Yes	Yes	Yes	Yes
Controls for establishment characteristics?	Yes	Yes	Yes	Yes	Yes
R^2	0.270	0.521	0.539	0.517	0.521
Number of Observations	7,578,426	2,437,528	171,787	58,707	10,156

	By Major Industry Group				
	Construction	Manufacturing	Retail Trade	Finance & Prof. Services	Local Services
$\ln D_j$	0.084 (0.019)	0.072 (0.016)	0.064 (0.016)	0.101 (0.012)	0.056 (0.005)
College Share, C_j	0.882 (0.214)	0.911 (0.135)	0.770 (0.099)	1.134 (0.124)	0.803 (0.099)
Year effects?	Yes	Yes	Yes	Yes	Yes
Controls for establishment characteristics?	Yes	Yes	Yes	Yes	Yes
R^2	0.154	0.279	0.254	0.219	0.280
Number of Observations	982,179	635,839	2,554,622	1,795,447	2,730,177

Notes: Table reports estimates from the regression of the log of average establishment earnings on the listed variables for our sample of establishment-year observations from the LBD. Establishment characteristics include the log of establishment employment, a dummy for whether the establishment is part of a multi-unit firm, fixed effects for age, and fixed effects for four-digit SIC. Standard errors, clustered by CBSA, are in parentheses.

Table 5. Establishment-Level Relations between Earnings and Density, Instrumental Variables Regressions

	All Establishments			
	(1)		(2)	
	OLS	IV	OLS	IV
$\ln D_j$	0.127	0.138	0.098	0.100
	(0.006)	(0.019)	(0.007)	(0.020)
College Share, C_j			0.898	1.588
			(0.099)	(0.279)
Year effects?	Yes	Yes	Yes	Yes
Controls for establishment characteristics?	No	No	Yes	Yes
R^2	0.017	0.010	0.317	0.315
Number of Observations	7,761,264			

Notes: Table reports estimates from the regression of the log of average establishment earnings on the listed variables for our sample of establishment-year observations from the LBD for the 283 CBSAs for which geological and climate data are available. Establishment characteristics include the log of establishment employment, a dummy for whether the establishment is part of a multi-unit firm, fixed effects for age, and fixed effects for four-digit SIC. Instruments include the fraction of the CBSA that is water-covered, the fraction above 1000m elevation, an index of the ruggedness of the land, the average annual temperature and moisture, the number of growing days, and the fraction of the land containing a set of 8 different soil types. Standard errors, clustered by CBSA, are in parentheses.

Table 6. Establishment-Level Relations between Earnings and Density across the Earnings Distribution

	Lowest Quintile	Second Quintile	Middle Quintile	Fourth Quintile	Highest Quintile
I. Within-Quintile Regression of Earnings on Density, Unconditional					
$\ln D_j$	0.080	0.083	0.096	0.110	0.144
	(0.011)	(0.009)	(0.008)	(0.007)	(0.008)
R^2	0.022	0.212	0.294	0.332	0.086
II. Within-Quintile Regression of Earnings on Density, Controlling for College Share					
$\ln D_j$	0.067	0.064	0.072	0.084	0.117
	(0.011)	(0.008)	(0.007)	(0.007)	(0.009)
College Share, C_j	0.637	0.919	1.103	1.213	1.285
	(0.119)	(0.101)	(0.107)	(0.119)	(0.145)
R^2	0.023	0.270	0.381	0.422	0.102
III. Within-Quintile Regression of Earnings on Density, Controlling for College Share and Characteristics					
$\ln D_j$	0.067	0.063	0.071	0.083	0.102
	(0.012)	(0.008)	(0.007)	(0.007)	(0.007)
College Share, C_j	0.640	0.913	1.089	1.188	1.116
	(0.107)	(0.099)	(0.104)	(0.116)	(0.133)
Controls for establishment characteristics?	Yes	Yes	Yes	Yes	Yes
R^2	0.104	0.296	0.407	0.446	0.278
Number of Observations	2,034,039	2,051,268	2,057,161	2,057,383	2,056,753

Notes: Table reports estimates from the regression of the log of average establishment earnings on log 1990 population density, and 1990 college share, where listed, within each quintile an establishment-year observation's CBSA-specific earnings distribution. All regressions include a year dummy. Establishment characteristics include the log of establishment employment, a dummy for whether the establishment is part of a multi-unit firm, fixed effects for age, and fixed effects for four-digit SIC. Standard errors, clustered by CBSA, are in parentheses.

Table 7. Statistics on the Earnings Distribution of Surviving Entrants in High- and Low-Density CBSAs

	Entrants after Surviving Five Years			
Statistic	**Low-Density CBSAs**	**High-Density CBSAs**	***High-Low Density Difference***	***High-Low Difference, Relative to Entry Values***
Mean Earnings				
Median Earnings				
Standard Deviation				
Interquartile Range				
90-10 Ratio				
50-10 Ratio				
Entrant Survival Rate	0.485	0.479		
Observations	38,820	332,018		

Notes: Table reports distributional statistics of earnings of entering establishments that survived to their fifth year, pooled across the top quarter (high density) or bottom quarter (low density) of CBSAs, ranked by 1990 population density. Statistics are based on an estimate of average establishment earnings that controls for establishment characteristics (the log of establishment employment, a dummy for whether the establishment is part of a multi-unit firm, fixed effects for age, and fixed effects for four-digit SIC).

Table 8. Summary Statistics of Relocating Establishments

Mean	**Relocating Establishments**	**Non-Relocating Establishments**	**Difference**
(log) Earnings			
Size (employees)	19.67	19.08	0.59 (0.30)
Age (years)	6.34	8.59	-2.25 (0.02)
Percent in Multi-Unit Firms	23.5	27.3	-3.8 (0.1)
(log) Density at Origin	5.947	5.936	0.011 (0.004)
(log) Density at Destination	5.730	---	
College Share at Origin	22.78	22.39	0.39 (0.02)
College Share at Destination	22.07	---	
Observations	81,766	7,799,688	

Notes: Table reports summary statistics for establishments who relocated from one CBSA to another versus establishments that remained in place. Standard errors for the difference between statistics are in parentheses.

Table 9. Earnings-Density Relationship Estimates based on Establishment Relocations

	Continuous Establishments			
	Levels Specification		**First Difference Specification**	
	(1)	(2)	(1)	(2)
(*change in*) $\ln D_j$	0.101 (0.007)	0.072 (0.008)	0.008 (0.004)	0.007 (0.003)
(*change in*) *College Share,* C_j		0.915 (0.091)		0.008 (0.030)
Year effects?	Yes	Yes	Yes	Yes
Controls for establishment characteristics?	No	Yes	No	Yes
R^2	0.017	0.354	0.001	0.167
Number of Observations	7,881,354			

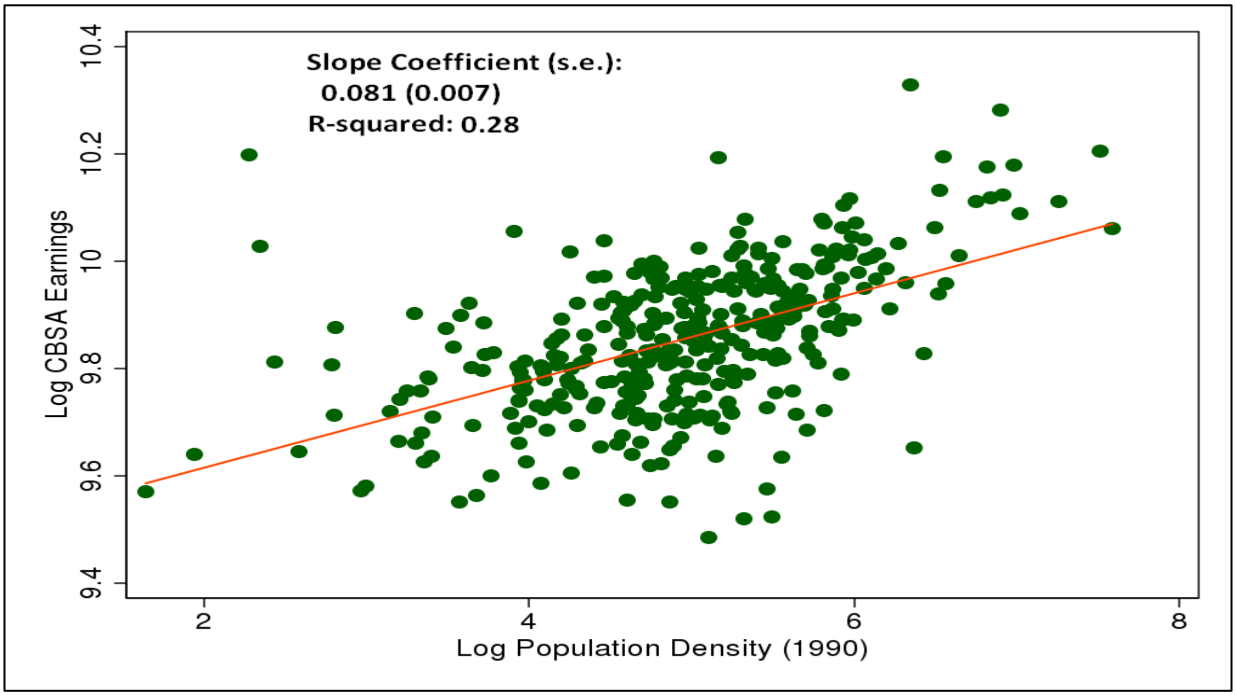
Notes: Table reports estimates from the regression of the log of average establishment earnings on the listed variables, or the change in (log) earnings on the change in the listed variables, for our sample of establishment-year observations from the LBD. Establishment characteristics in the levels specification include the log of establishment employment, a dummy for whether the establishment is part of a multi-unit firm, fixed effects for age, and fixed effects for four-digit SIC. Establishment characteristics in the first-difference specification include the change in establishment employment and age fixed effects. First differences are measured year-over-year (1991-92 or 1996-97 for continuous establishments observed in 1992 or 1997, respectively). Standard errors, clustered by CBSA, are in parentheses.

Table 10. Statistics on the Earnings Distribution of Entrants in High- and Low-Density CBSAs

All Entrants				
Statistic	**Low-Density CBSAs**	**High-Density CBSAs**	***High-Low Density Difference***	***High-Low Difference, Relative to Incumbents***
Mean Earnings	9.640	9.852	0.212	0.013
Median Earnings	9.756	9.936	0.180	0.000
Standard Deviation	0.883	0.910	0.026	-0.014
Interquartile Range	0.976	0.997	0.022	0.002
90-10 Ratio	2.104	2.179	0.075	-0.028
Observations	80,092	693,139		
Entrants of Multi-Unit Firms				
Statistic	**Low-Density CBSAs**	**High-Density CBSAs**	***High-Low Density Difference***	***High-Low Difference, Relative to Incumbents***
Mean Earnings	9.860	10.023	0.163	-0.049
Median Earnings	9.906	10.045	0.139	-0.041
Standard Deviation	0.677	0.737	0.060	0.033
Interquartile Range	0.639	0.701	0.062	0.040
90-10 Ratio	1.408	1.554	0.146	0.071
Observations	18,530	165,691		

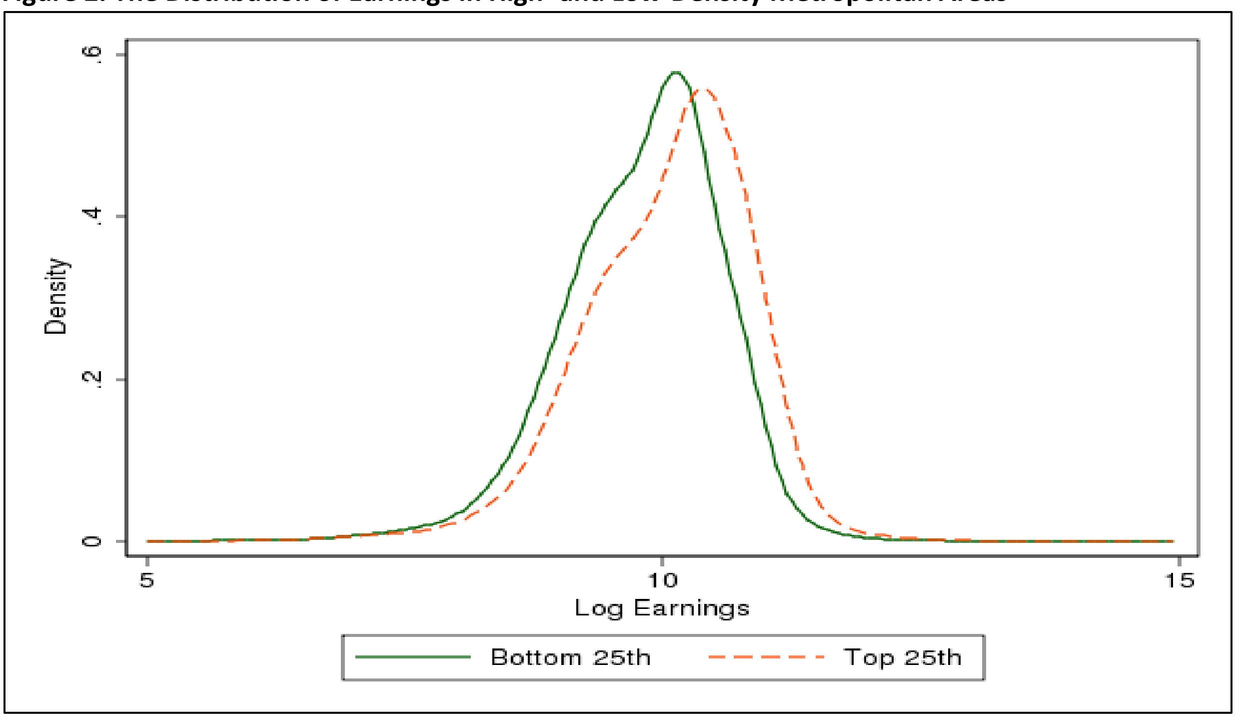
Notes: Table reports distributional statistics of earnings of entering establishments pooled across the top quarter (high density) or bottom quarter (low density) of CBSAs, ranked by 1990 population density. Statistics based on an estimate of average establishment earnings that controls for establishment characteristics (the log of establishment employment, a dummy for whether the establishment is part of a multi-unit firm, fixed effects for age, and fixed effects for four-digit SIC).

Figure 1. The Relationship between CBSA Earnings and Density



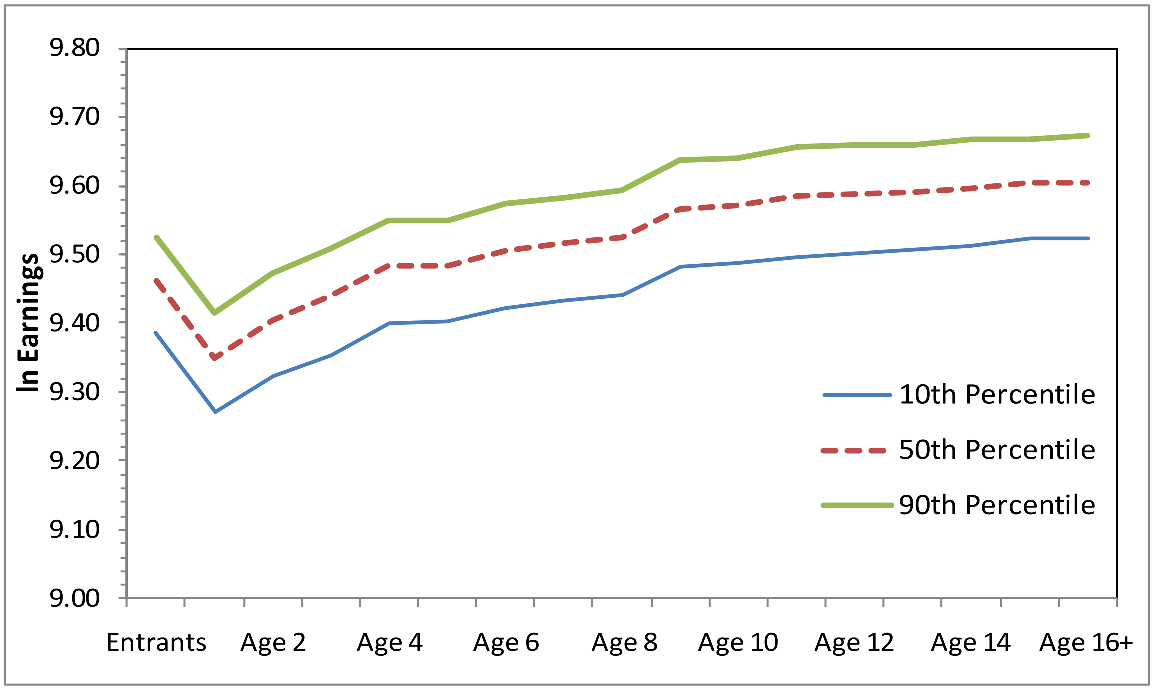
Note: The figure plots the relation between the log of average earnings on log 1990 population density for the 363 CBSAs of our sample, along with the fitted linear trend and its slope and *R*-squared value.

Figure 2. The Distribution of Earnings in High- and Low-Density Metropolitan Areas



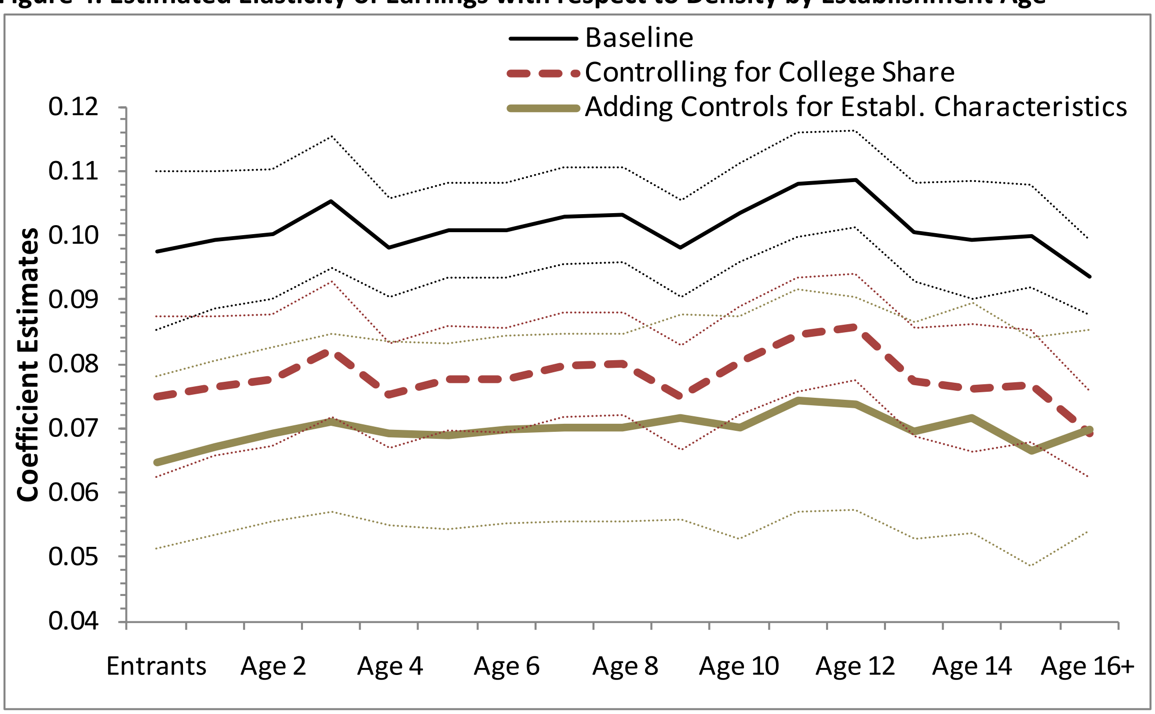
Note: The figure plots the kernel density estimates of the distribution of log average earnings for the 91 CBSAs in the bottom quartile of the density distribution (solid line) and the 91 CBSAs in the top quartile of the density distribution (dashed line), respectively.

Figure 3. Estimated Elasticity of Earnings with respect to Density by Establishment Age



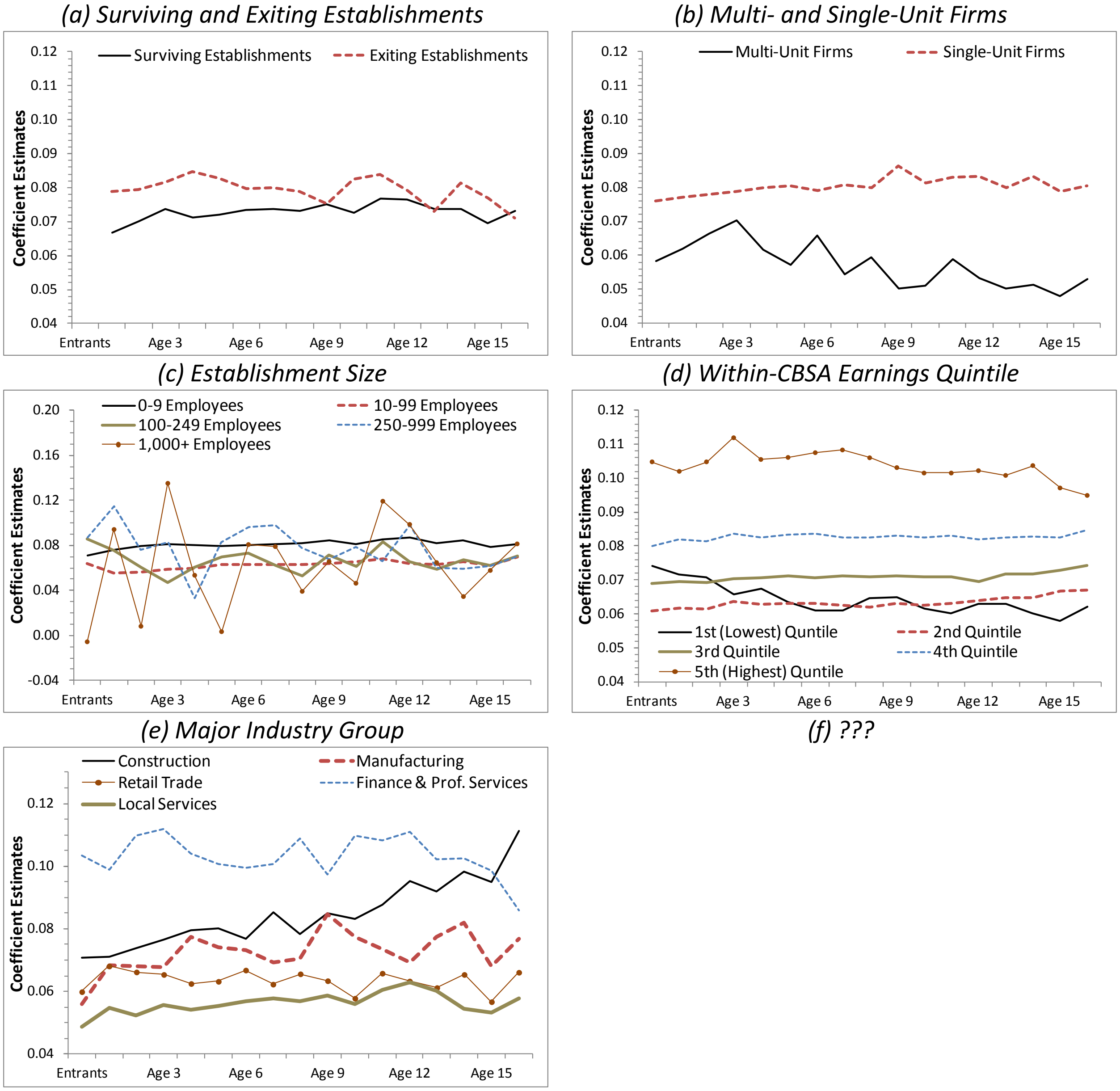
Note: The figure plots predicted earnings from the estimation of equation (3) in the text for CBSAs at the 10th, 50th, and 90th percentiles of the population density distribution. See text for estimation details.

Figure 4. Estimated Elasticity of Earnings with respect to Density by Establishment Age



Note: The figure plots the predicted elasticity of earnings with respect to density as a function of age. Estimates come from equation (3) in the text. See text for details. Thin dashed lines represent standard error bands, with standard errors clustered by CBSA.

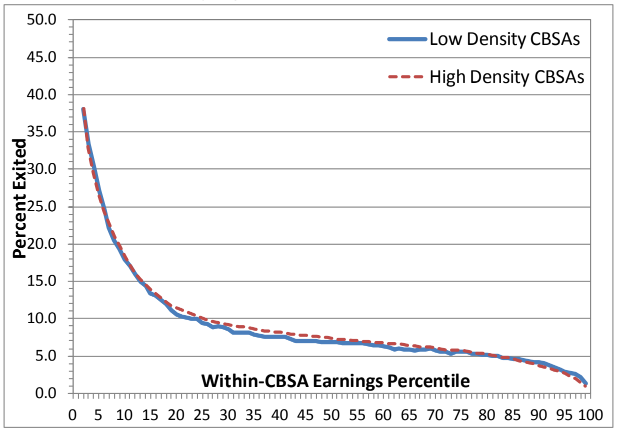
Figure 5. Elasticity of Earnings with respect to Density by Establishment Age and Sub-Group



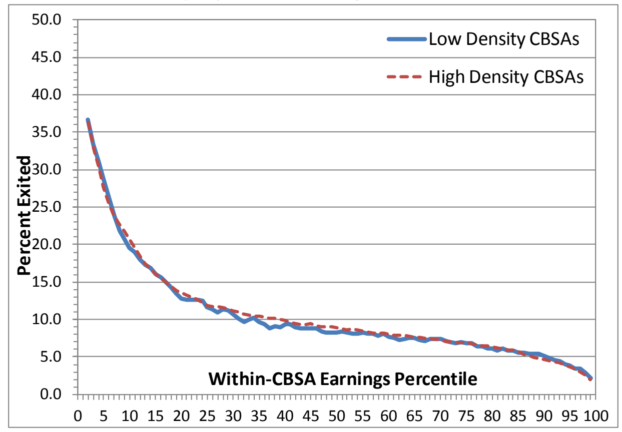
Note: Each panel of the figure plots the predicted elasticity of earnings with respect to density as a function of age. Estimates come from equation (3) in the text. See text for details. Thin dashed lines represent standard error bands, with standard errors clustered by CBSA.

Figure 7. Establishment Exit Probabilities by CBSA Earnings Percentile

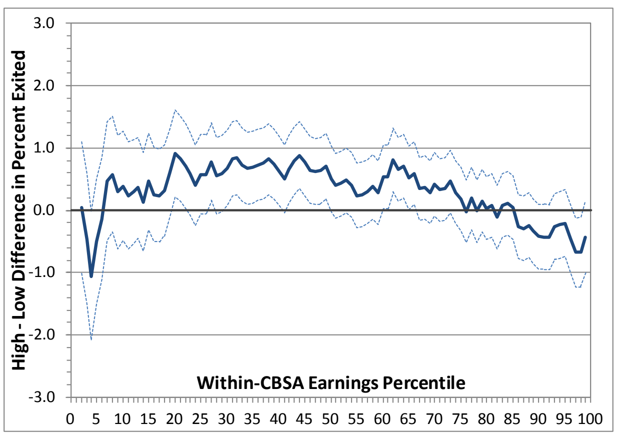
(a) Probability of Exit: All Establishments



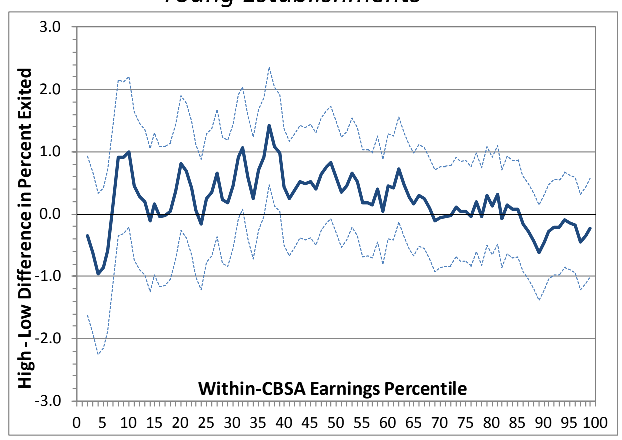
(b) Probability of Exit: Young Establishments



(c) High – Low Difference: Exit Probability of all Establishments

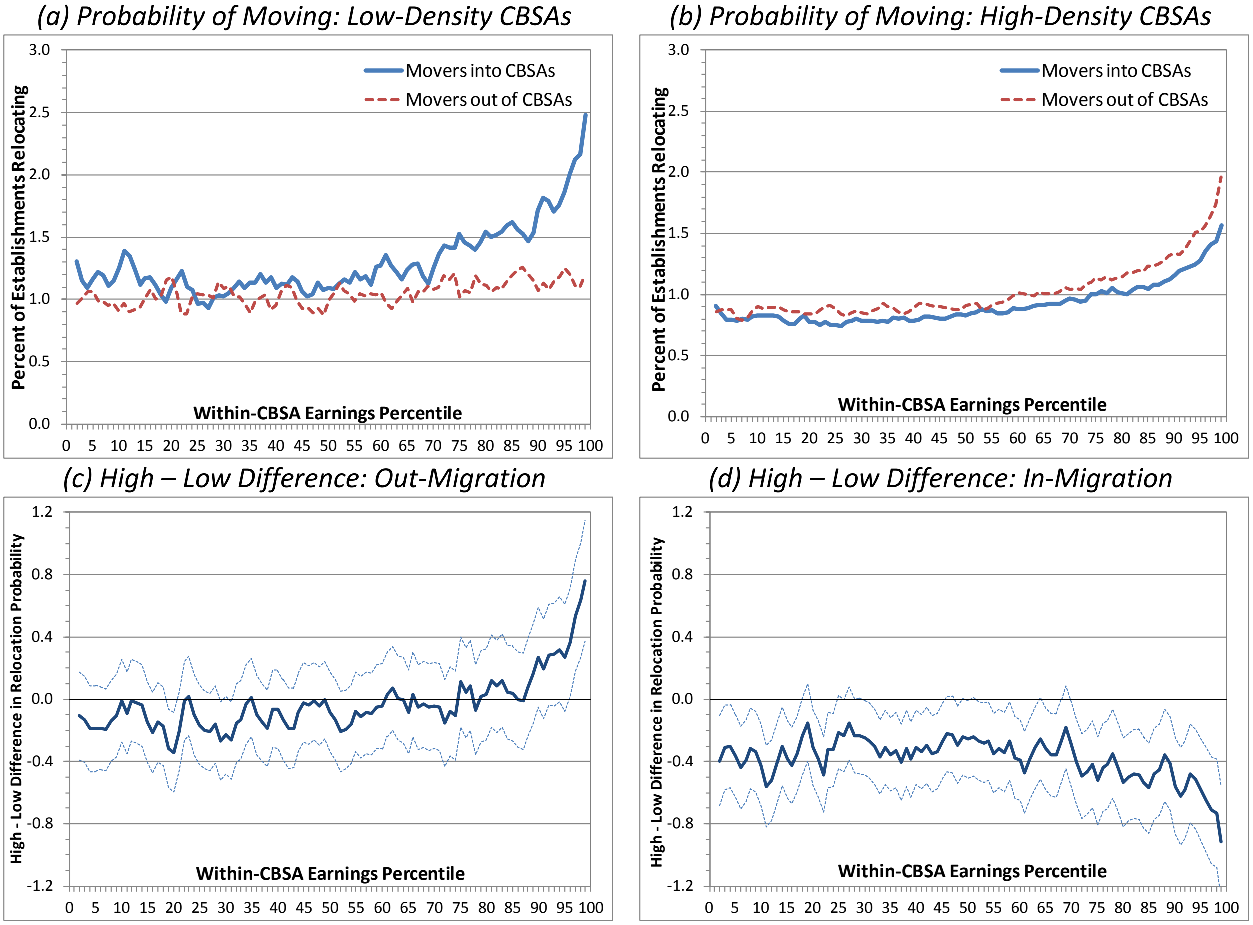


(d) High – Low Difference: Exit Probability of Young Establishments



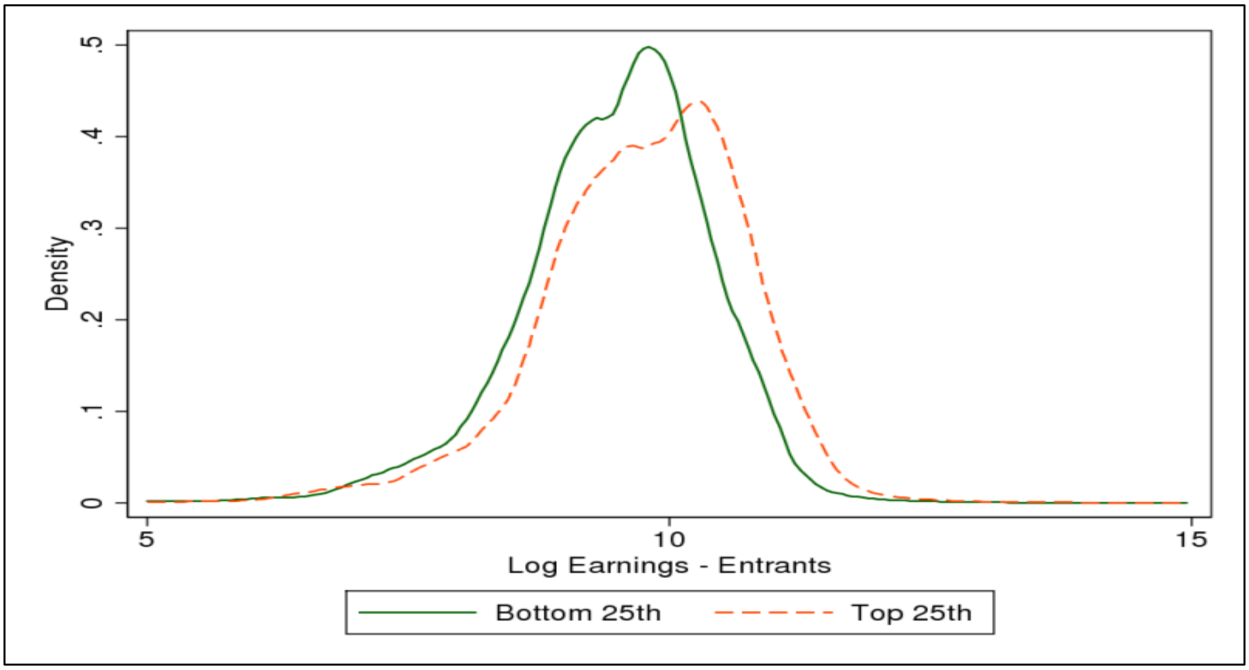
Note: Top panels report the probability of an establishment exiting by percentiles of the within-CBSA earnings distribution for CBSAs grouped into the highest and lowest quartiles of the CBSA density distribution. Bottom panels report the difference in exit probabilities between high-density and low-density CBSAs. The left panels are for all establishments while the right panels are for establishments aged 5 years or less. All probabilities and earnings are conditional on establishment characteristics (size, age, industry, multi-unit firm status). All panels report 3-percentile, centered averages to smooth the estimates.

Figure 8. Establishment Relocation Probabilities by CBSA Earnings Percentile



Note: Top panels report the probability of an establishment relocating from one CBSA to another by percentiles of the within-CBSA earnings distribution for CBSAs grouped into the highest and lowest quartiles of the CBSA density distribution. Bottom panels report the difference in out-migration and in-migration, respectively between high-density and low-density CBSAs. All probabilities and earnings are conditional on establishment characteristics (size, age, industry, multi-unit firm status). All panels report 3-percentile, centered averages to smooth the estimates.

Figure 9 . Distribution of Earnings of Entering Establishments by Urban Density



Note: The figure plots the kernel density estimates of the distribution of log average earnings for entering establishments in CBSAs in the bottom (solid line) and top (dashed line) quartiles of the density distribution, respectively. The top panel shows the unconditional distributions, while the bottom panel shows the distributions of earnings relative to mean CBSA earnings.