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

September 2011

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 estimate an elasticity of average establishment earnings with respect to population density of about 10 percent. This estimate falls to just over 7 percent once we control for establishment characteristics and the share of the local population with a college degree. Looking across the distribution of establishments, we find that establishments in the highest quintile of the earnings distribution exhibit a density premium that is almost double that of the lowest quintile. We also find that the density premium does not change over an establishment's tenure within a city, which is true after a variety of controls and within all subgroups of the data. We show that both entrants and exits in dense cities tend to have higher earnings and a more disperse earnings distribution, but we caution against interpreting either result as evidence in support of a role for either the sorting of high-earnings establishments into dense cities or the selection of low-earnings establishments out of local markets, respectively.

Keywords: density premium, urban agglomeration, establishment entry and exit
JEL Codes:

** Correspondence: Faberman: [jfaberman @ frbchi.org](mailto:jfaberman@frbchi.org), Freedman: Freedman@cornell.edu. We thank Nate Baum-Snow 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.*

I IIII III I I I IIII I I

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 mirrors previous research on workers, focusing on the average earnings at a particular establishment. Examining average earnings at the establishment provides a direct comparison to previous findings on workers' earnings density premium. 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.

We first estimate the density premium separately at the city and establishment levels and show that the establishment data produce an establishment premium that is similar to the worker-based

---

[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 than the earnings premium.

studies. As Figure 1 shows, we find that the unconditional 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 remains above 6 percent when controls for the share of the population with a college degree and establishment characteristics are added.[3]

We then examine the potential role of heterogeneity. We test whether the density premium varies with the average earnings of an establishment and find that the density premium rises with average establishment earnings. Controls for establishment characteristics and the education of the local population do not qualitatively affect the result. Unconditionally, establishments in the highest quintile of the average earnings density exhibit a density premium that is nearly double that of establishments in the lowest quintile, 8.0 percent versus 14.4 percent. Even after the controls are added, there exists a sizeable and significant difference, 6.7 percent versus 10.2 percent. This may suggest that high-earnings (or arguably, high-productivity) establishments reap differentially greater returns to urban density relative to low-earnings establishments. This is analogous to similar findings for workers (e.g., Glaeser and Mare, 2001; Glaeser and Resseger, 2010; and Gould, 2007) that show evidence of "skill-biased" returns to agglomeration. That is, workers with higher observable skills tend to gain a higher earnings premium for working in dense locations. In this sense, one can interpret our findings as evidence of "productivity-biased" returns to agglomeration.

Next, we test whether the density premium rises as establishments age, which would suggest the existence of returns to density that accumulate with the city-specific tenure of an establishment. Since establishments rarely relocate, their age and city tenure are essentially synonymous. Potentially increasing returns to city tenure may accrue through faster learning or greater knowledge spillovers

[3] The difference between the city-level and establishment-level estimates reflect the fact that establishments are implicitly weighted by their employment in the city-level estimate.

(that are accumulated over time) in denser metropolitan areas. We find, however, that as establishments age, their estimated density premium remains essentially constant. 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; de la Roca and Puga, 2010). The independent findings may appear as a paradox, since establishments are a collection of workers, but we illustrate that changes in the composition of establishment's workforce, such as a trend toward replacing older, high-tenure workers with younger, low-tenure workers, can reconcile the findings, as can the presence of *decreasing* returns to an establishment's tenure within a dense city. The latter would be consistent with the nursery city mechanism described by Duranton and Puga (2001). As we illustrate, however, establishment data alone cannot separately identify potential worker-specific and establishment-specific returns to city tenure.

A concern with estimating the returns to urban density for workers is that highly-skilled workers may self select into dense cities, which would imply that any positive relation between earnings and density is at least partly driven by sorting. Several recent models of urban growth have emphasized a role for a similar sorting mechanism for firms; i.e., high-productivity firms may disproportionate enter dense locations. Recent models have also stressed a potential role a selection effect via firm exit.[4] In other words, denser areas may also be more competitive, leading to more exit and a greater truncation of the firm productivity distribution from below.

Dealing with the sorting issue for establishments is difficult because they rarely relocate, making it practically impossible to distinguish sorting from a city-specific return to density that affects productivity. Our finding that the density premium is constant over the life an establishment further complicates matters, since it negates the opportunity to use across-city variation within age cohorts as a

[4] See, for example, Combes et al. (2009), Behrens, Duranton, and Robert-Nicoud (2010), and Holmes, Hsu, and Lee (2010).

potential identification channel. In light of this, we present evidence on the distribution of average earnings of entrants in high-density and low-density metropolitan areas. Our aim is to establish some basic facts on the differences between entrants in these locations to guide future research on the role of firm sorting. In doing so, we distinguish between entrants that are single-unit firms and entering establishments that are part of a multi-unit firm. The working hypothesis is that the former group is more likely to locate where their entrepreneurs live, and are thus relatively less likely to choose their location across a range of cities.

Overall, we find that entrants have lower earnings, on average, and wider dispersion of their earnings distribution, relative to incumbents. Further, entrants in the highest quartile of metropolitan areas, ranked by their population density, have earnings that are 26 percent higher than entrants in the lower quartile, with an interquartile range of their earnings distribution that is 9 percent wider and a 90-10 ratio of the distribution that is 19 percent larger, with much of the differences in dispersion concentrated among entrants in the top half of the earnings distribution. Controlling for differences in establishment characteristics only reduces these estimates slightly, but a difference-in-difference calculation of the entrant differences relative to the differences between incumbent establishments nearly no disparity between the two groups of entrants. Comparing the difference between entrants of multi-unit firms yields similar results. Further, when we calculate a difference-in-difference estimate of the differences between the multi-unit entrants relative to the differences between all entrants, we find that entrants of multi-unit firms in high-density cities have about 4 percent lower relative earnings, on average, and no real difference in the dispersion of their earnings distribution. On its surface, we view the latter results as inconsistent with a sorting story, though further research is warranted.

To gauge the prevalence of selection, we perform a similar exercise for the relative earnings of exiting and surviving establishments. As with entrants, we find that exiting establishments in the densest quartile of metropolitan areas have higher earnings and a greater amount of dispersion in their earnings

distribution. The distribution of surviving establishments, however, does not exhibit much of the behavior consistent with selection, as suggested by standard models of agglomeration and selection (e.g., Syverson, 2004). Survivors have higher earnings, on average, but also have greater dispersion in their earnings distribution, rather than evidence of truncation. An examination of exit probabilities, however, suggests that while establishments are less likely to exit dense areas overall, those with high average earnings are relatively more likely to exit these areas. Overall, we interpret these results as at least suggestive of a role for firm selection in explaining the observed density premium for earnings though, again, further research is needed.

I Ⅲ I II I

We use establishment data from the Longitudinal Business Database (LBD) of the Census Bureau.[5] 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 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.

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

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.[6] 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 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.

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,

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

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 often consider the average earnings of an establishment as a proxy for its productivity. In the absence of an output measure to check the validity of this assumption, we assess it by replicating the results of Syverson (2004) using our average earnings measure in lieu of Total Factor Productivity (TFP). The Syverson study is a useful benchmark because, like our study, it focuses on the relationship between establishment-level dynamics and urban density, and consequently, is interested in similar moments of the data. 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.

Our replication uses LBD data for the same subsample as Syverson: 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, iii) and 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. We report the results of our exercise, along with the original results from Syverson (2004), in Table 1. We use an OLS regression specification that includes the log of construction employment density and year dummies on the right-hand side (Syverson's "Model 2"). As one can see, we find quantitatively similar results for all six moments. If anything, we find even stronger results than Syverson, though this may be a consequence of excluding non-metropolitan areas from the analysis. 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).[7] 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.

[illegible]

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).[8] These estimates are roughly in line with estimates previously found using data on individuals rather than establishments.[9]

Figure 1 reports the CBSA-level elasticity of wages with respect to density. We also estimate the elasticity using establishment-level data. This has the advantage of allowing 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 also lets us estimate the density premium for different subsets of the data. 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 consequently include a year dummy, and cluster our standard errors by CBSA. For comparison to the aggregate results, we also examine the

[7] 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.
[8] The coefficient on the college share is 0.33 with a standard error of 0.10.
[9] See, for example, Ciccone and Hall (1996) and Sveikauskas (1975).

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, and 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).

The results for the full sample of establishment-years observations results appear in Table 3. Unconditionally, we find a somewhat higher density premium at the establishment level relative to the aggregate level, at 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.[10] In each 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.

The composition of establishments differs across cities. It is plausible that the above estimates mask wide heterogeneity in the density premium across different subgroups of the data. For example, smaller establishments may experience a different density premium than larger establishments. The same may be true of younger versus older establishments. In Table 4, we re-estimate (2) separately for different subgroups of the data.

The first groups we examine are entering and exiting establishments.[11] 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 are significantly different from their counterpart in Table 3. We

---

[10] This is consistent with recent work by Lehmer and Möeller (2010), who find an urban density premium that persists after controlling for firm size.

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

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.[12] We also estimate the density premium for five establishment size classes and for young (1-5 years), middle-aged (5-10 years), and older (11 or more years) establishments. In both cases, we find no significant differences in the earnings premium across groups in either category.

[illegible]

One goal of this paper is to document evidence of the urban density premium experienced by firms comparable to the existing evidence of the density premium experienced by workers. Two recent findings that stand out in the research on workers are that the density premium is higher for more skilled workers (e.g., Glaeser and Mare, 2001; Gould, 2007; Baum-Snow and Pavan, 2010b; and Glaeser and Resseger, 2010) and that workers city tenure-wage profiles tend to rise more steeply in more dense cities (e.g., Glaeser and Mare, 2001; de la Roca and Puga, 2010). One can interpret the first finding as evidence of “skill-biased” returns to agglomeration. That is, workers with higher observable skills tend to gain a higher earnings premium for working in dense locations. Our first exercise examines whether there is a similar “productivity-biased” return to agglomeration for establishments, in the sense of a higher premium for establishments with higher average earnings. One can interpret the second finding as evidence of greater human capital gains that accumulate the longer an individual remains in a dense city. Such gains may stem from a greater degree of knowledge spillovers or a faster degree of “learning.” Our second exercise examines whether firms exhibit similar increasing returns to agglomeration the longer they remain in a dense city.

[12] 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.

[illegible]

Figure 2 shows the full distribution of earnings for two subsets of the data: establishments in the top quartile and bottom quartile of CBSAs, ranked by their population density.[13] The data are pooled over both years and the kernel density estimates use an unconditional earnings measure. The figure clearly shows a rightward shift of the entire earnings distribution for establishments in the high-density CBSAs. The earnings distribution in the densest CBSAs also exhibits greater dispersion. Given the heterogeneity in average establishment earnings across metropolitan areas, and given the observed differences across the full distribution of earnings in Figure 2, it is natural to ask whether the establishments experience a greater density premium at different points of the earnings distribution.

We study this question by disaggregating the earnings distribution for each CBSA into quintiles and estimating the earnings premium for establishments within each quintile separately. That is, we first identify which quintile of their CBSA-specific earnings distribution each establishment belongs to. We then group all establishments into five separate groups based on their CBSA-specific quintile. Finally, we estimate the regression in (2) for each group. Thus, a higher estimated coefficient on density for one group implies that establishments within that quintile of the distribution *for their own CBSA* have a higher density premium.

Table 5 presents the results. We estimate three specifications. The top rows of Table 5 present the results when regressing the log of average establishment earnings on the log of density and controls for year only. The middle rows of Table 5 present the results of adding in college share as an additional control. The bottom rows of Table 5 present the results of the full specification in (2), which additionally controls for establishment characteristics (size, age, multi-unit firm status, industry).[14]

---

[13] 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.

[14] 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 estimates show a clear monotonic rise in the estimated density premium with average establishment earnings across the distribution. 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) almost double that of the lowest quintile (8.0 percent). One can interpret this finding as implying that establishments with higher average earnings (or arguably, with higher productivity) exhibit a greater return to urban density, or as evidence of what we would refer to as "productivity-biased" returns to agglomeration. Controlling for college share and establishment characteristics lowers the estimated elasticity for all quintiles and decreases the disparity between the highest and lowest quintiles to 3.5 percentage points, but the monotonic relation between the density premium and average establishment earnings remains.

[illegible]

We next explore whether the earnings density premium for establishments is dynamic; i.e., whether the premium rises with an establishment's tenure at a particular location. For establishments, the appropriate metric to measure their tenure in a metropolitan area is their age. Unlike workers, firms rarely migrate across areas. Furthermore, what little evidence exists on the propensity of establishments to relocate across locations suggests that its frequency is almost an order of magnitude smaller than the frequency that establishments enter and exit the economy.[15] Thus, for nearly all establishments, their city tenure equals their age.

To test for a rising premium with tenure for establishments in more dense areas, we repeat our establishment-level regressions of (log) earnings on population density including fixed effects for the age

---

the within-CBSA ranking of their resulting *residual* earnings measure. The approach produced very similar results to the ones reported in Table 4.

[15] For example, Duranton and Puga (2001) report that 4.7 percent of their sample of French establishments moved across regions during a three-year period. In contrast, Table 2 shows that over 10 percent of establishments enter the economy, on average, *per year*. Evidence on job creation and job destruction rates in France and the U.S. (Abowd, Corbel, and Kramarz, 1999; Spletzer et al., 2004) suggest that the two metrics are comparable across countries.

of the establishment and an interaction between these fixed effects and density. As with the earlier analyses, we also run two additional specifications that control for the CBSA college share and for both the college share and establishment characteristics. The general form of the regression we run is

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

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.[16] 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 urban 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 age for each specification 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

[16] 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.

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 four panels of the figure report the coefficients $\hat{\beta}^1 + \hat{\zeta}(a)$ from the estimation of the full specification in (3) separately for continuing and exiting establishments, for establishments in multi-unit and single-unit firms, for establishments by size class, and for establishments by their within-CBSA earnings quintile, respectively. 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. 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.

[illegible]

The results of Figures 3 and 4 are striking in their contrast to the analogous results for workers found by Glaeser and Mare (2001), Baum-Snow and Pavan (2010b), de la Roca and Puga (2010), and others. It could imply one of several things. For instance, it might imply that the dynamic benefits of urban density are captured primarily by workers rather than firms. This would suggest that factors such as human capital accumulation and moving up a job ladder play an important role. It might also imply

that all the returns to density for firms are realized at entry at a particular location. This would an extreme version of the finding of Glaeser and Mare (2001), who find that much of the earnings growth realized by workers is at the time of in-migration to a more dense area. Such a finding might imply that firms tend to benefit more from fixed locational benefits, such as local comparative advantage in specialized inputs or demand, rather than benefits that accumulate over time, such as knowledge spillovers.

In both cases, however, one must grapple with the fact the establishments of a city are a collection of the same workers whose earnings exhibit steeper wage-tenure profiles in more dense cities. Aggregation alone suggests that more must be going on. Changes in the skill and city tenure composition of workers within the establishment over time could help reconcile the joint findings of a steeper tenure profile with density for workers and a flat tenure profile with density for establishments. Given the findings in Figures 3 and 4, accounting for such compositional changes using only establishment data only allows one, in the best case, only to place an upper bound on the establishment-level returns to urban density.

To see this, consider the following Mincer-style earnings regression for the i^{th} worker:

(4) $$\ln w_{iejt} = \alpha_i + \alpha_e + \beta \ln D_j + \lambda n_{iet} + \eta_0 \tau_{ijt} + \varphi_w a_{et} + \eta(\tau) \ln D_j + \zeta_w(a) \ln D_j + \varepsilon_{iejt}.$$

The regression is a function of fixed worker and firm characteristics (α_i,and α_e), density, worker tenure at both the establishment, n_{iet}, and the city, τ_{ijt}, establishment age, and the interactions between worker's city tenure and density and the establishment's age and density. For exposition, we ignore the controls for college share and establishment characteristics and let the age effects enter linearly. For simplicity, take the geometric mean of the workers' log earnings to represent log earnings at the establishment level. Let a bar over the variable denote the establishment level average of that variable at a point in time. The establishment-level analog to (4) is then,

(5) $$\ln w_{ejt} = \bar{\alpha}_{et} + \alpha_e + \beta \ln D_j + \lambda \bar{n}_{et} + \eta_0 \bar{\tau}_{et} + \varphi_w a_{et} + \overline{\eta(\tau)}_{et} \ln D_j + \zeta_w(a) \ln D_j + \epsilon_{ejt}(a).$$

Comparing (5) to (3), it is clear that the estimate of $\hat{\zeta}(a)$ obtained from the regression of (3) will capture the effects of both $\overline{\eta(\tau)}_{et}$ and $\zeta_w(a)$, since both will change as the establishment ages. It will also include much of the average worker characteristics terms in the regression's error term. One might think of taking first differences of (5), but because the composition of an establishment's workforce can change, this does little to improve estimation and identification of $\zeta_w(a)$:

(6) $$d \ln w_{ejt} = \Delta\bar{\alpha}_{et} + \lambda \Delta\bar{n}_{et} + \eta_0 \Delta\, \bar{\tau}_{et} + \varphi_w + \Delta\overline{\eta(\tau)}_{et} \ln D_j + [\zeta_w(a) - \zeta_w(a-1)] \ln D_j + \Delta\epsilon_{ejt}.$$

Furthermore, even in the case where the establishment employs the same workers from one period to the next (which is indistinguishable from an establishment that replaces its turnover with no net employment change in the establishment data) the first difference will still include the interaction term between urban density and worker tenure. Thus, it would still be difficult to identify any establishment-specific change in the density premium with age. This will be especially true if the relationship between the worker's wage-tenure profile and density is nonlinear (i.e., $\partial^2\eta(\tau)/\partial\tau^2 \neq 0$).

If one were to take the findings on increasing returns to city tenure for workers as given, it would imply that $\partial\eta(\tau)/\partial\tau > 0$. If turnover at the establishment level is not sufficiently high, it would in turn imply that $\Delta\overline{\eta(\tau)}_{et} > 0$, on average. Given our empirical result of $\hat{\zeta}(a) \approx 0$ for all a, these results would together imply that, at best, the establishment-level density premium is indeed flat with respect to city tenure and may even be decreasing with tenure. The latter result would be particularly interesting since it would appear consistent with the nursery city theory of Duranton and Puga (2001). In their model, firms enter denser cities to reap the benefits of agglomeration, but as they age and grow, the congestion costs of operating in a dense city eventually outweigh the returns to agglomeration, leading them to relocate to less dense locations.

Ultimately, further research using matched employer-employee data is needed to obtain a better identification of the true relationship between city tenure and the density premium for establishments. We note, though, that for the U.S. at least, the matched employer-employee data has its own shortcomings for an urban study. For example, it is not always possible to identify the location of the individual establishments of a firm. Further, establishment age and worker tenure can only be measured back to the start of the data, which is in the early 1990s for most U.S. states.

[illegible]

[illegible]

An issue for the interpreting our estimated density premium as a true return to urban density is sorting. That is, the potential exists that establishments in dense cities have higher earnings, on average, not because of city-specific factors but because more productive establishments tend to locate in denser cities. Similar concerns for sorting exist when estimating the returns to density for workers (i.e., high-skill workers may self-select into denser cities). Researchers often account for worker sorting by following their migration patterns and estimating the within-city density premium for migrating individuals. This approach is not feasible for establishments because they rarely relocate. Instead, the entry margin is the main mechanism that establishments use to locate in a new city. This makes it impossible to distinguish between a high-earnings entrant who endogenously chose to locate in a dense city and an entrant who is able to pay high average earnings because of location-specific returns they reap from an exogenously given city. Our findings on the age density interaction suggest that identification of the establishment-level return to density is even more difficult because any establishment-specific returns may be realized at entry, precisely when sorting is hardest to examine,

and may be constant over the life of the establishment, eliminating the ability to exploit variation across establishment cohorts as a source of identification.

Given these difficulties, we take the approach of presenting some new facts on the distribution of earnings for entering establishments. Our goal is to present new evidence to motivate future research on the role of sorting by firms. 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 between entrants of multi-unit and single-unit firms. The working hypothesis is that entrants of multi-unit firms are much more likely to choose the location of a new establishment from a set of metropolitan areas, whereas a new, single-establishment, firm that is more likely to start up where its entrepreneur is currently located.

a(s)7(i)-de(e)-8( )]TJ 0.002 T10027Tw 8.1543exsiestaiesna3ain (g)3(ts)]T4 0.001 4 T-020522-38.359 -2.43(e)-36(s

W31 (g)3(( )]T7 0.0 5 1nes

table also quantifies the differences in dispersion between the two groups. The standard deviation of earnings across entrants is 1.04 in the high-density CBSAS versus 0.99 in the low-density CBSAs. Their interquartile range is also 9 percent larger, while the 90-10 ratio 19 percent larger. The bottom panel of the table shows that controlling for differences in establishment characteristics across the entrants in the two groups of CBSAs only marginally changes the key findings: entrants in the high-density CBSAs now have 21.2 percent higher earnings, the difference in their interquartile range is 2.2 percent, and the difference in their 90-10 ratio is now 7.5 percent. As before, most of the differences in earnings are among entrants in the top half of the earnings distribution.

When we calculate the difference-in-difference estimate of these statistics for entrants relative to incumbents, however, it implies that the relative differences in the earnings distributions of entrants are small. After controlling for establishment characteristics, entrants in high-density CBSAs have earnings that are only 1.3 percent higher, in relative terms, exhibit essentially the same interquartile range as entrants in low-density CBSAs, and actually have a somewhat (2.8 percent) smaller 90-10 ratio, driven mostly by relatively higher earnings among entrants at the 10th percentile of the distribution.[17]

Table 7 shows similar findings for geographic differences between the entrants of multi-unit firms. Entrants of multi-unit firms within high-density CBSAs have 22.5 percent higher earnings, an interquartile range that is 9.7 percent wider, and a 90-10 earnings ratio that is 21 percent higher. When one controls for differences in establishment characteristics earnings for these entrants remain 16.3 percent higher in the high-density CBSAs and the dispersion of earnings remains greater as well, with an interquartile range that is 6.2 percent wider and a 90-10 earnings ratio that is 14.6 percent higher.

Given our working hypothesis that the entrants of multi-unit firms are more likely to sort into locations than the entering single-unit firms, we would expect a sorting mechanism to generate

[17] Most of these differences remain statistically significant, since the standard error of mean earnings across these subgroups ranges between 0.1 and 0.5 percentage points, but their economic significance is arguably small.

differences in the relative distributions of these two groups. We calculate another difference-in-difference estimate, this time subtracting the difference in the earnings distributions of all entrants in high-density versus low-density CBSAs from the differences in the earnings distributions of entrants from multi-units firms in high-density versus low-density CBSAs. The results are in the last column Table 7. If anything, we find that the entrants of multi-unit firms have *lower* relative earnings, on average, though they still maintain a somewhat more disperse earnings distribution. The latter is driven primarily by differences between the bottom halves of the distributions. While it is difficult to conclude much from these differences without an underlying model of firm entry decision-making, given our hypothesis on the location decision-making of multi-unit firms, we would consider this evidence as inconsistent with a strong role for sorting.

[illegible]

Establishments can also affect the estimated density premium through their exit behavior. If dense cities are locations where competition among firms is relatively strong, establishments with a lower level of productivity may be more likely to exit, all else equal. This selection effect on the exit margin would generate a positive relationship between average earnings and density through a lower-truncation of the earnings distribution. Combes et al. (2009) examine the role of selection, and while they find that a selection effect is evident in the data, they attribute little of the observed density premium to the selection effect—they find that differences in selection across cities, as measured from differences in firms' productivity distribution, are small. 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). We tend to agree that an examination of the selection effect requires a separate examination of exiting and surviving establishments.

We proceed with an analogous exercise as before, except that we now compare exits in both absolute and relative terms. We start with some basic facts about the earnings distribution of exits. Figure 8 presents the analog to Figure 6 for exiting establishments in our sample. As one can see, exiting establishments have considerably lower average earnings, on average, and a substantially wider distribution of average earnings relative to survivors. Figure 9, the analog to Figure 7, shows that there are also clear differences between the exiting establishments in high-density and low-density CBSAs. The earnings distribution for exiting establishments in high-density CBSAs is shifted to the right, relative to exits in low-density CBSAs, and exhibits somewhat greater dispersion.

Table 8 quantifies these differences in the distribution. Exiting establishments in the high-density CBSAs have earnings that are 26.5 percent higher, on average, than exiting establishments in the low-density CBSAs. Their earnings distribution is also more disperse, with a difference between their interquartile ranges of 3.9 percent and a difference between their 90-10 ratios of 9.4 percent. On the surface, this appears to be consistent with a selection mechanism by which high-earnings establishments are more likely to exit in denser areas. Differences between the earnings distributions of surviving establishments presents more mixed evidence. Survivors have earnings that are 23.9 percent higher, on average, in high-density CBSAs, but their earnings distribution is also more disperse, with an interquartile range that is 6.4 percent wider and a 90-10 ratio that is 9.8 percent. The greater dispersion is at odds with the implications of standard models of firm selection, which postulate that the selection mechanism should generate truncation of the earnings or productivity distribution from below (and hence a *less* disperse distribution).

Finally, we examine how the probability an establishment exits relates to urban density. *A priori*, it is not clear whether, conditional on other characteristics, the exit rate should be higher in more dense areas. What should be true is that, conditional on other characteristics, establishments with higher earnings should be more likely to exit denser metropolitan areas. We test for this by using a linear

probability regression where the dependent variable is a dummy for whether the establishment exited. In our full specification, as before, we regress this on the log of population density, the college share, and establishment characteristics (size, age, multi-unit firm status, industry), including a dummy for year. We also include the log of average establishment earnings and an interaction term between density and earnings. The latter term is the variable of interest, since a positive coefficient would indicate that high-earnings establishments are more likely to exit in denser metropolitan areas

Our results are in Table 9. If we regress the exit probability on only density and average earnings, a doubling of density increases the probability of exit by about 1 percentage point. Controlling for college share reduces this estimate only slightly. If we include the interaction between average earnings and density, then density alone decreases the probability of exit, but that probability rises with the establishment's average earnings. The effect is slightly stronger if we control for establishment characteristics. We interpret this as evidence that selection may play a role in generating the observed density premium for establishments.

Overall, we take the evidence on the earnings distributions and exit patterns of establishments as mixed with respect to the role of selection. Exiting establishments clearly have higher earnings in more dense areas, but there is no evidence that this is because of a truncation of their earnings distribution from below (i.e., a selection effect). Further work, potentially focusing on the differences in earnings and outcomes among young establishments between high-density and low-density areas, may provide stronger evidence of any potential role for selection in generating a positive density premium.

I III I I I II I II I I I

In this paper, we present new evidence on the earnings premium associated with urban density for establishments. We do so in order to present new stylized facts on the relation between average

establishment earnings, which we consider a proxy for productivity, and urban density that parallel similar facts other studies have found for the earnings of workers.

In general, we find an elasticity of average establishment earnings with respect to density that is similar to the estimates obtained in earlier studies for workers. Furthermore, we find that this elasticity is increasing in average establishment earnings. Establishments in the highest quintile of the earnings distribution have an estimated elasticity that is nearly double that of establishments in the lowest quintile, even after controlling for differences in establishment characteristics. This finding of “productivity-biased” returns to agglomeration is also similar to some recent studies of workers that find evidence of a “skill-biased” return to agglomeration. We also find that the estimated density premium does not change over the lifetime of the establishment, suggesting that longer tenure in a city does little to increase any returns to agglomeration establishments may experience. This is at odds with models where firms accumulate dynamic returns to locating in a dense city through faster learning or greater knowledge spillovers, but may be consistent with the nursery city mechanism suggested by Duranton and Puga (2001). It also contrasts with the findings of steeper wage profiles with respect to city tenure for workers in denser areas found by several recent studies. We argue, however, that with establishment data alone, one cannot identify the true change in the density premium over an establishment’s life because of unobserved changes in workforce composition.

We note that the sorting of high-earnings establishments into dense metropolitan areas could account for some of our estimated density premium, but we concede that two issues make the identification of such sorting in the data difficult. The first is the fact that establishments rarely relocate, eliminating within-establishment differences in city density as a potential source of identification. The second is our finding that the density premium does not change over the establishment life cycle, eliminating within-cohort differences in the density premium as a potential source of identification. Nevertheless, we examine differences among entrants in high-density and low-density metropolitan

areas to establish some stylized facts to guide future research. We find that entrants in high-density cities have are higher and have a more disperse distribution of earnings in absolute terms, but that these differences disappear when compared relative to the differences among incumbents. Comparing the earnings of entrants of multi-unit firms to the earnings of all entrants, we find that the former group actually has lower relative earnings. Given that we expect entrants of multi-unit firms to be more likely to endogenously choose their location, we argue that this evidence does not support a strong role for sorting in accounting for a positive density premium, though further research is needed on the topic.

Finally, we perform a similar analysis on exiting establishments to see whether firm selection (via exit) out of a metropolitan area may account for part of the estimated density premium. We find that exiting establishments have higher earnings in denser metropolitan areas and that the probability of that a high-earnings establishment exits is more likely in a dense area. At the same time, we find that the earnings distribution of surviving establishments has a higher mean in denser areas, but shows no evidence of truncation, as a standard model of firm selection would imply. Overall, we consider our findings to be mixed evidence at best for a role for firm selection. Again, further research is needed on the issue.

I I II I I I I I I

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.

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

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.

Our evaluation method depends on whether the establishment has positive employment in the preceding and proceeding years. If an establishment has positive employment in both years, we impute its average earnings in year *t* if

i) $|g_{eit}^{w}| > 1.96$ and $|g_{ei,t+1}^{w}| > 1.96$, or

ii) $|\widetilde{w}_{ei,t}^{t}| > 1.96$.

In other words, we impute the value if the change in average earnings if it is within 4 percent of the maximum absolute value of the growth rate measure, or if the value relative to the industry median is within 4 percent of its maximum absolute value using our relative earnings measure. The value that we choose for imputed average earnings, $\widehat{w}_{eit}$, depends on whether the leading and lagging average earnings estimates exhibit extreme values:

$$\widehat{w}_{eit} = \begin{cases} \frac{1}{2}(w_{ei,t-1} + w_{ei,t+1}) & \text{if } \left| \dfrac{\frac{1}{2}(w_{ei,t+1}+w_{ei,t-1})-\overline{w}_{it}}{\frac{1}{2}\left[\frac{1}{2}(w_{ei,t+1}+w_{ei,t-1})+\overline{w}_{it}\right]} \right| \leq 1.96, \text{ and} \\ \overline{w}_{it} & \text{otherwise.} \end{cases}$$

In words, if the average of average earnings in t + 1 and t – 1, relative to median industry earnings is not an extreme value (as defined by a modified version of our relative earnings measure $\widetilde{w}_{ei,t}^{t}$), we use the average of the leading and lagging average earnings estimates. Otherwise, we use the median industry earnings.

If an establishment only has positive employment in *t* – 1, we impute the value of average earnings in *t* if

i) $|g_{eit}^{w}| > 1.96$, or

ii) $|\widetilde{w}_{ei,t}^{t}| > 1.96$.

The value we use in the imputation depends on whether the average earnings measure in *t* – 1 is an extreme value, relative to median industry earnings:

$$\widehat{w}_{eit} = \begin{cases} w_{ei,t-1} & \text{if } \left|\widetilde{w}_{ei,t-1}^{t}\right| \leq 1.96, \text{ and} \\ \overline{w}_{it} & \text{otherwise.} \end{cases}$$

Similarly, if an establishment only has positive employment in *t* + 1, we impute the value of average earnings in *t* if

i) $\left|g_{ei,t+1}^{w}\right| > 1.96$, or

ii) $\left|\widetilde{w}_{ei,t}^{t}\right| > 1.96$.

The value we use in the imputation depends on whether the average earnings measure in *t* + 1 is an extreme value, relative to median industry earnings:

$$\widehat{w}_{eit} = \begin{cases} w_{ei,t+1} & \text{if } \left|\widetilde{w}_{ei,t+1}^{t}\right| \leq 1.96, \text{ and} \\ \overline{w}_{it} & \text{otherwise.} \end{cases}$$

Finally, for establishments that only have positive employment in year *t*, we only impute the average earnings value if $\left|\widetilde{w}_{ei,t}^{t}\right| > 1.96$, in which case we use $\overline{w}_{it}$was the imputed value.

Combined, all four cases of the imputation algorithm affect relatively few establishments, with less than one-tenth of one percent of establishments affected in any given year in our sample.

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

	Estimate of Demand Density Elasticity	
Moment	**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	0.079	0.080	0.058	
	(0.011)	(0.013)	(0.010)	(0.009)	
College Share, C_j	1.135	0.815	0.945	0.700	
	(0.129)	(0.100)	(0.108)	(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	0.064	0.067	0.075	0.071
	(0.010)	(0.010)	(0.009)	(0.012)	(0.013)
College Share, C_j	0.926	0.769	0.862	0.749	0.703
	(0.104)	(0.084)	(0.111)	(0.113)	(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 Age Group				
	1 to 5 Years	6 to 10 Years	11+ Years		
$\ln D_j$	0.077	0.072	0.071		
	(0.012)	(0.009)	(0.008)		
College Share, C_j	0.841	0.927	0.903		
	(0.104)	(0.095)	(0.090)		
Year effects?	Yes	Yes	Yes		
Controls for establishment characteristics?	Yes	Yes	Yes		
R^2	0.292	0.319	0.339		
Number of Observations	3,740,749	2,058,852	3,393,264		

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

(a) Unconditional Earnings

Statistic	*Incumbents* Low-Density CBSAs	*Incumbents* High-Density CBSAs	*Entrants* Low-Density CBSAs	*Entrants* High-Density CBSAs	*Difference, Entrants*	*Difference-in-Difference*
Mean Earnings	9.651	9.890	9.376	9.636	0.260	0.021
Standard Deviation	0.795	0.837	0.988	1.035	0.047	0.005
Skewness	0.006	0.064	-0.240	-0.149	0.090	0.032
Interquartile Range	1.005	1.068	1.230	1.321	0.090	0.027
90th Percentile	10.583	10.892	10.534	10.878	0.344	0.035
50th Percentile	9.674	9.905	9.438	9.680	0.242	0.011
10th Percentile	8.653	8.864	8.141	8.295	0.154	-0.057
90-10 Ratio	1.930	2.028	2.393	2.583	0.190	0.092

(b) Earnings Conditional on Establishment Characteristics

Statistic	*Incumbents* Low-Density CBSAs	*Incumbents* High-Density CBSAs	*Entrants* Low-Density CBSAs	*Entrants* High-Density CBSAs	*Difference, Entrants*	*Difference-in-Difference*
Mean Earnings	9.653	9.852	9.640	9.852	0.212	0.013
Standard Deviation	0.652	0.692	0.883	0.910	0.026	-0.014
Skewness	-0.272	-0.127	-0.517	-0.431	0.085	-0.060
Interquartile Range	0.745	0.765	0.976	0.997	0.022	0.002
90th Percentile	10.372	10.639	10.582	10.856	0.274	0.007
50th Percentile	9.685	9.865	9.756	9.936	0.180	0.000
10th Percentile	8.873	9.037	8.478	8.677	0.199	0.035
90-10 Ratio	1.499	1.602	2.104	2.179	0.075	-0.028

Notes: Table reports distributional statistics of earnings for all establishment-year observations pooled across the top quarter (high density) or bottom quarter (low density) of CBSAs, ranked by 1990 population density. The top panel reports the statistics based on the raw estimate of average establishment earnings, while the bottom panel reports the statistics based on an estimate 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). Difference-in-difference estimates report the difference between high- and low- density earnings statistics relative to the difference between high- and low-density incumbent statistics.

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

(a) Unconditional Earnings

Statistic	*Multi-Unit Entrants* Low-Density CBSAs	*Multi-Unit Entrants* High-Density CBSAs	*Difference, MU Entrants*	*Difference- in-Difference, MU Entrants – Incumbents*	*Difference- in-Difference, Multi-Unit – All Entrants*
Mean Earnings	9.756	9.981	0.225	-0.014	-0.035
Standard Deviation	0.824	0.912	0.088	0.045	0.040
Skewness	-0.258	-0.151	0.107	0.049	0.017
Interquartile Range	1.084	1.181	0.097	0.034	0.007
90th Percentile	10.731	11.063	0.332	0.023	-0.012
50th Percentile	9.775	9.999	0.224	-0.007	-0.018
10th Percentile	8.807	8.929	0.122	-0.089	-0.032
90-10 Ratio	1.924	2.134	0.210	0.112	0.020

(b) Earnings Conditional on Establishment Characteristics

Statistic	*Multi-Unit Entrants* Low-Density CBSAs	*Multi-Unit Entrants* High-Density CBSAs	*Difference, MU Entrants*	*Difference- in-Difference, MU Entrants – Incumbents*	*Difference- in-Difference, Multi-Unit – All Entrants*
Mean Earnings	9.860	10.023	0.163	-0.036	-0.049
Standard Deviation	0.677	0.737	0.060	0.019	0.033
Skewness	-0.737	-0.503	0.234	0.089	0.149
Interquartile Range	0.639	0.701	0.062	0.042	0.040
90th Percentile	10.544	10.805	0.261	-0.006	-0.013
50th Percentile	9.906	10.045	0.139	-0.041	-0.041
10th Percentile	9.136	9.251	0.115	-0.049	-0.084
90-10 Ratio	1.408	1.554	0.146	0.043	0.071

Notes: Table reports distributional statistics of earnings for all establishment-year observations pooled across the top quarter (high density) or bottom quarter (low density) of CBSAs, ranked by 1990 population density. The top panel reports the statistics based on the raw estimate of average establishment earnings, while the bottom panel reports the statistics based on an estimate 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). Difference-in-difference estimates report the difference between high- and low- density earnings statistics relative to the difference between high- and low-density incumbent statistics.

Table 8. Statistics on the Earnings Distribution of Exits in High- and Low-Density CBSAs

(a) Unconditional Earnings

	Exits			
Statistic	**Low-Density CBSAs**	**High-Density CBSAs**	***Difference, Exits***	***Difference, Survivors***
Mean Earnings	8.822	9.087	0.265	0.239
Standard Deviation	1.082	1.126	0.044	0.043
Skewness	0.048	0.061	0.013	0.058
Interquartile Range	1.489	1.528	0.039	0.064
90th Percentile	10.196	10.514	0.318	0.309
50th Percentile	8.864	9.099	0.235	0.230
10th Percentile	7.400	7.624	0.224	0.211
90-10 Ratio	2.796	2.890	0.094	0.098

(b) Earnings Conditional on Establishment Characteristics

	Exits			
Statistic	**Low-Density CBSAs**	**High-Density CBSAs**	***Difference, Exits***	***Difference, Survivors***
Mean Earnings	9.058	9.267	0.209	0.198
Standard Deviation	0.969	1.001	0.032	0.043
Skewness	-0.173	-0.129	0.044	0.143
Interquartile Range	1.294	1.303	0.008	0.065
90th Percentile	10.201	10.445	0.244	0.271
50th Percentile	9.125	9.328	0.203	3.179
10th Percentile	7.796	7.96	0.164	0.161
90-10 Ratio	2.405	2.485	0.080	0.110

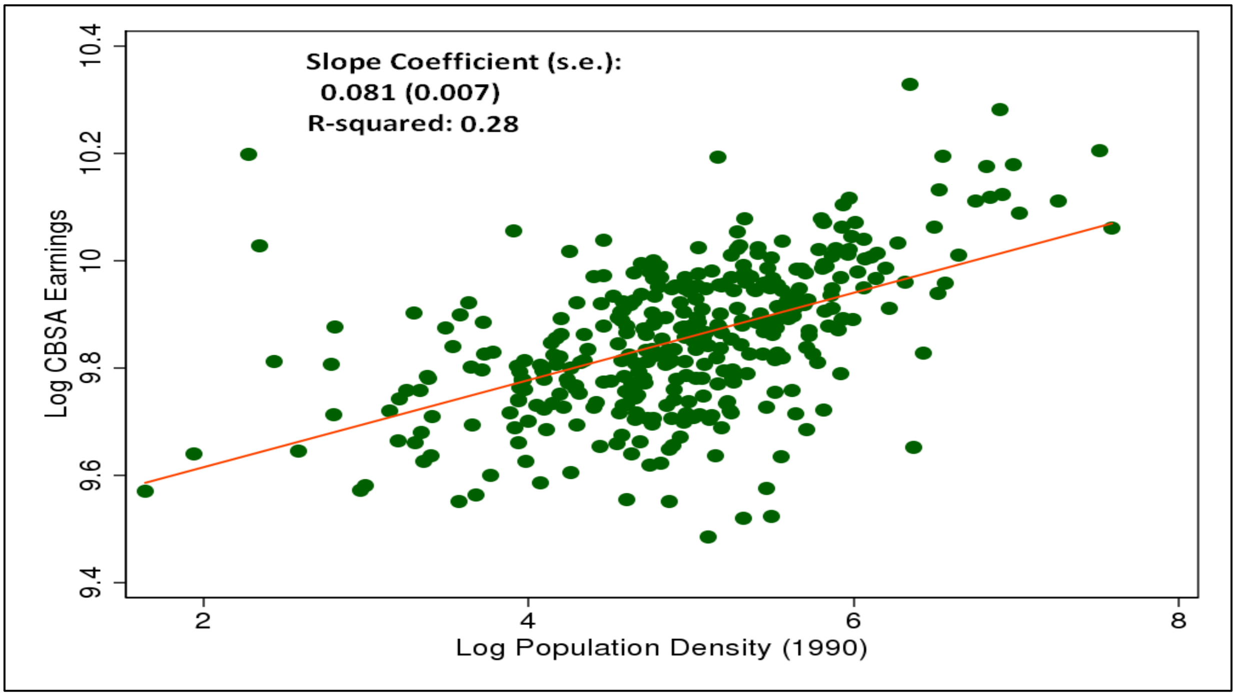
Notes: Table reports distributional statistics of earnings for all establishment observations pooled across the top quarter (high density) or bottom quarter (low density) of CBSAs, ranked by 1990 population density. The top panel reports the statistics based on the raw estimate of average establishment earnings, while the bottom panel reports the statistics based on an estimate 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). Difference-in-difference estimates report the difference between high- and low- density earnings statistics relative to the difference between high- and low-density incumbent statistics.

Table 9. The Relation between Survival Rates, Earnings, and Density

	(1)	(2)	(3)	(4)
$\ln D_j$	0.010	0.008	-0.022	-0.028
	(0.001)	(0.001)	(0.007)	(0.006)
$\ln w_{ejt}$	-0.080	-0.080	-0.099	-0.108
	(0.001)	(0.001)	(0.004)	(0.003)
$\ln D_j \times \ln w_{ejt}$			0.003	0.004
			(0.001)	(0.001)
College Share, C_j		0.100	0.100	0.097
		(0.016)	(0.016)	(0.013)
Year effects?	Yes	Yes	Yes	Yes
Controls for establishment characteristics?	No	No	No	Yes
R^2	0.00	0.00	0.06	0.18

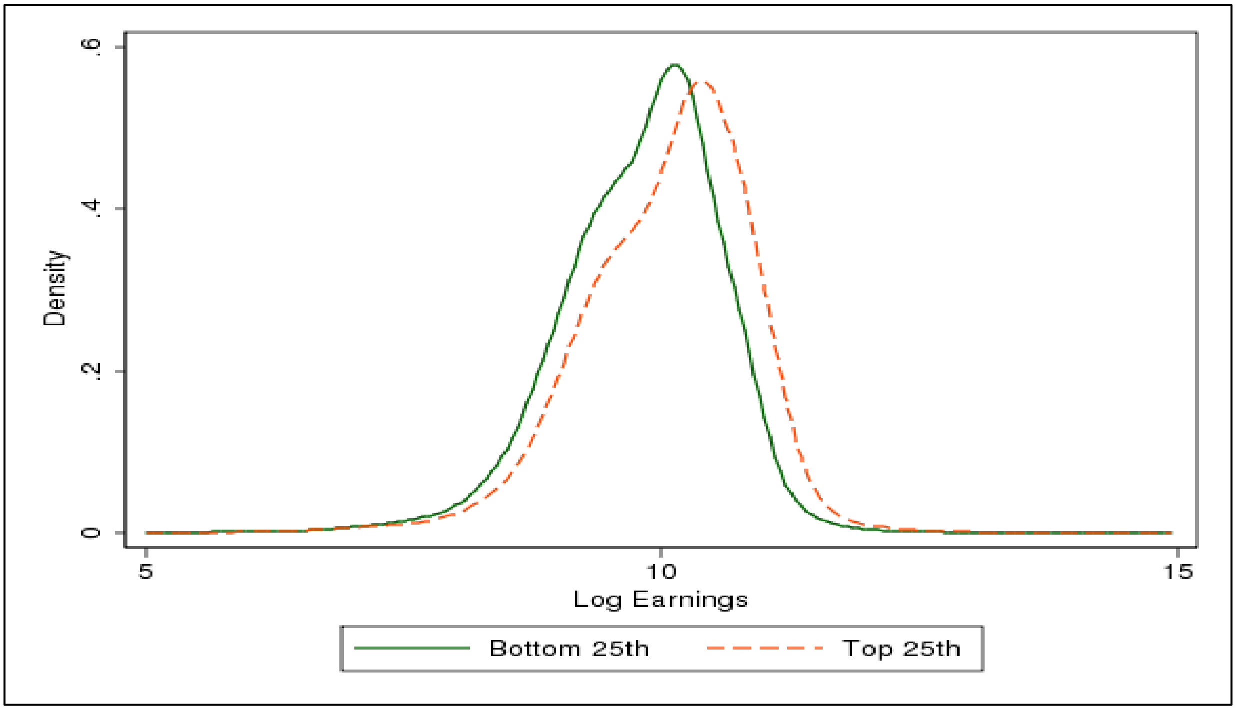
Notes: Table reports estimates from the regression of a dummy for whether an establishment exited on the listed variables for our sample of establishment-year observations from the LBD. *N* = 10,256,604. Establishment characteristics include the log of establishment employment, a dummy for whether the 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.

Figure 1. The Relationship between CBSA Earnings and Density



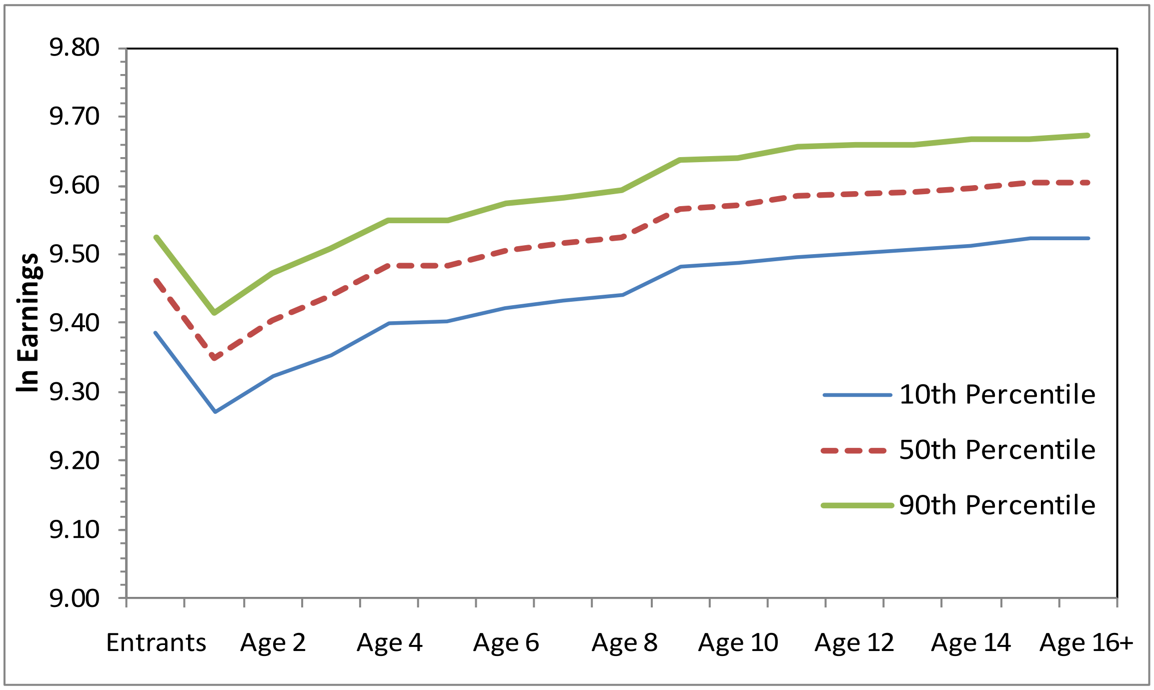
Note: The figure plots the relation between the log of average earnings on log 1990 population density for the 363 CBSAs of our sample, along with the fitted linear trend and its slope and *R*-squared value.

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



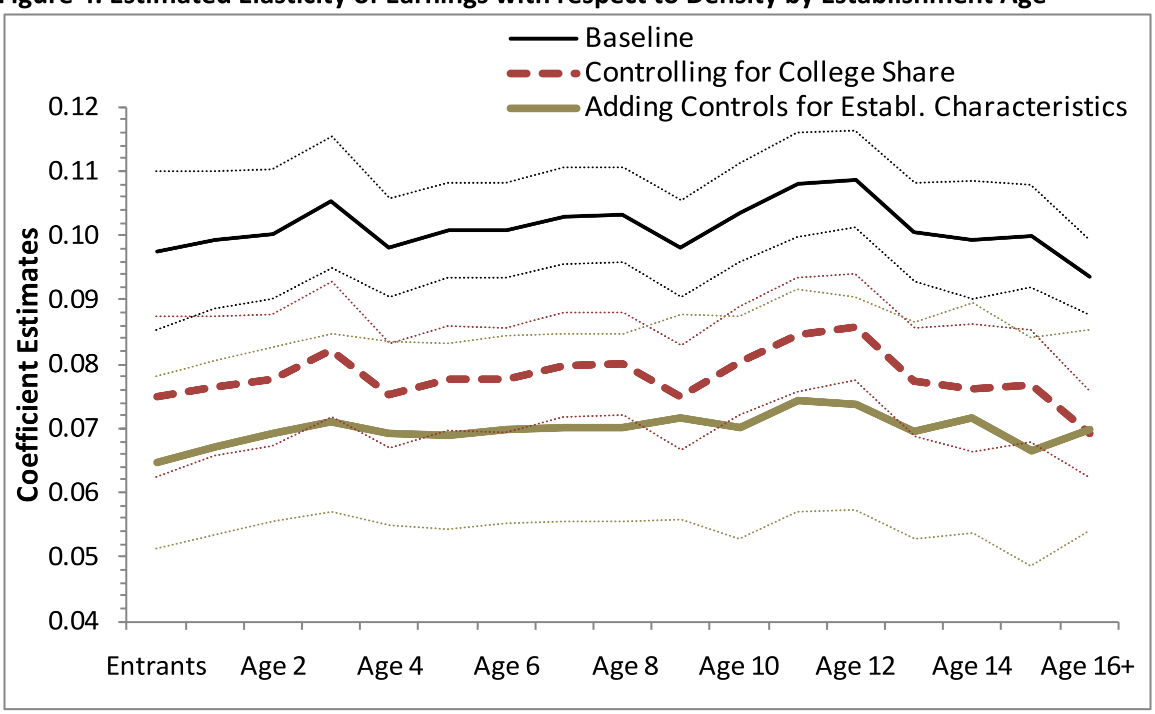
Note: The figure plots the kernel density estimates of the distribution of log average earnings for the 91 CBSAs in the bottom quartile of the density distribution (solid line) and the 91 CBSAs in the top quartile of the density distribution (dashed line), respectively.

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

(a) Surviving and Exiting Establishments

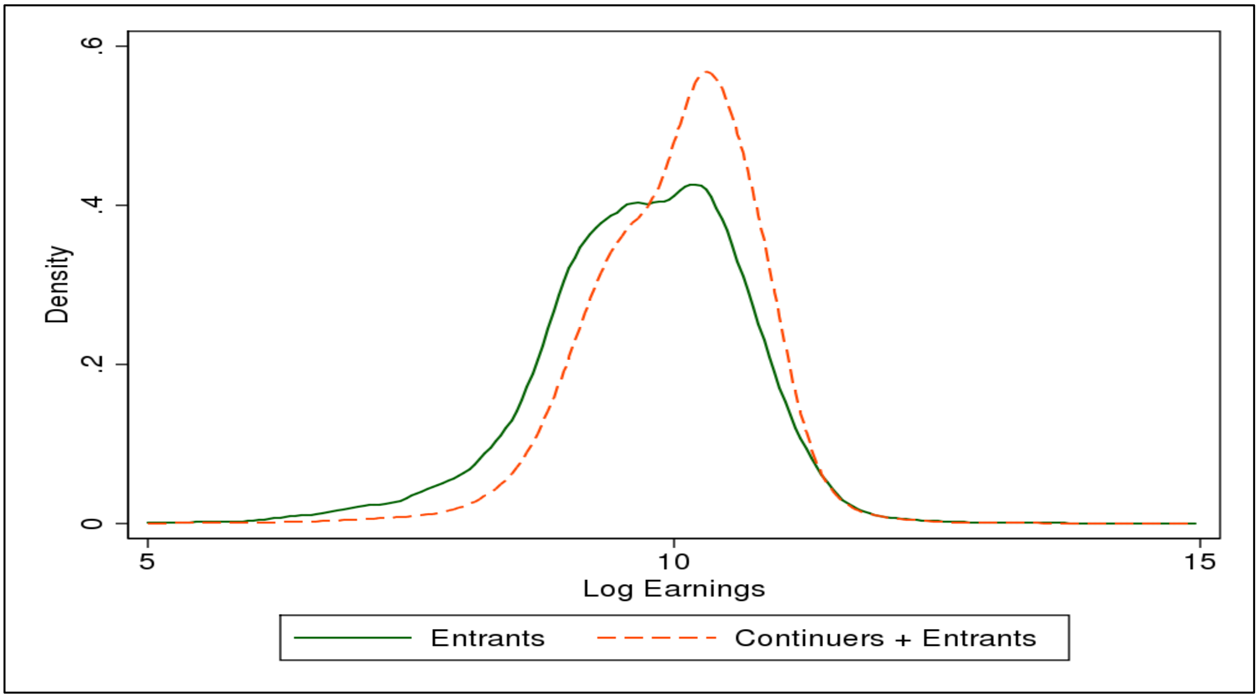
(b) Multi- and Single-Unit Firms

(c) Establishment Size

(d) Within-CBSA Earnings Quintile

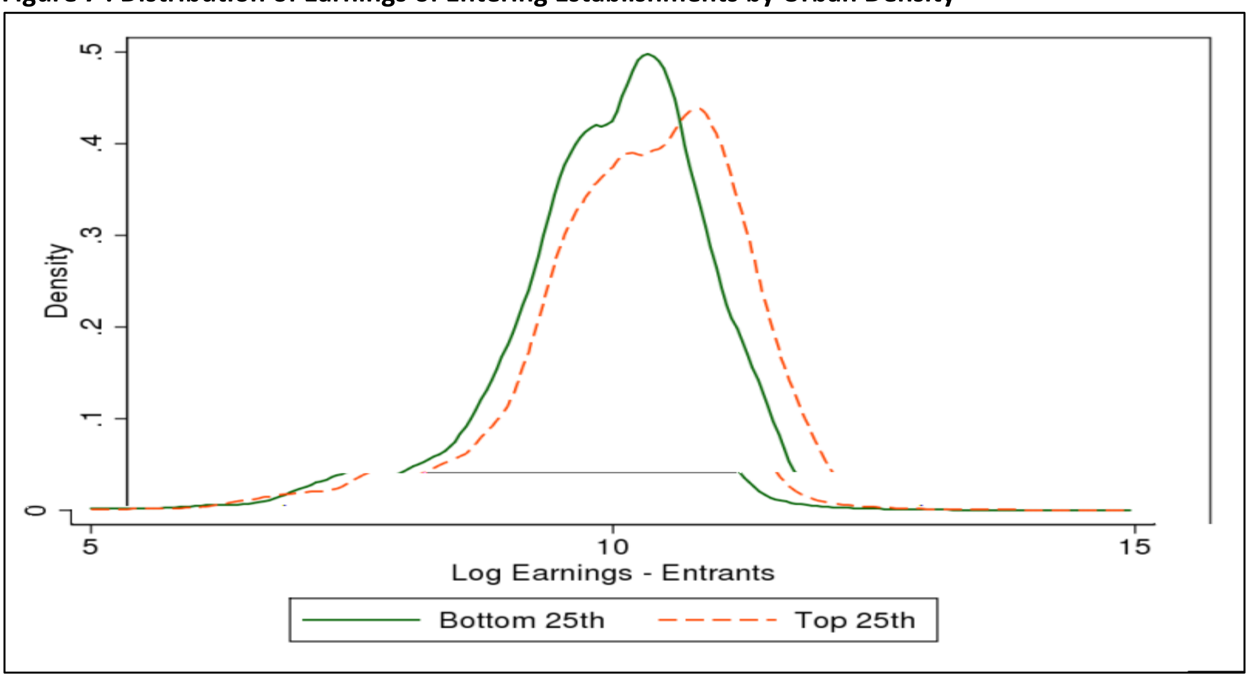
Note:

Figure 6. The Distribution of Earnings for Entering and Incumbent Establishments



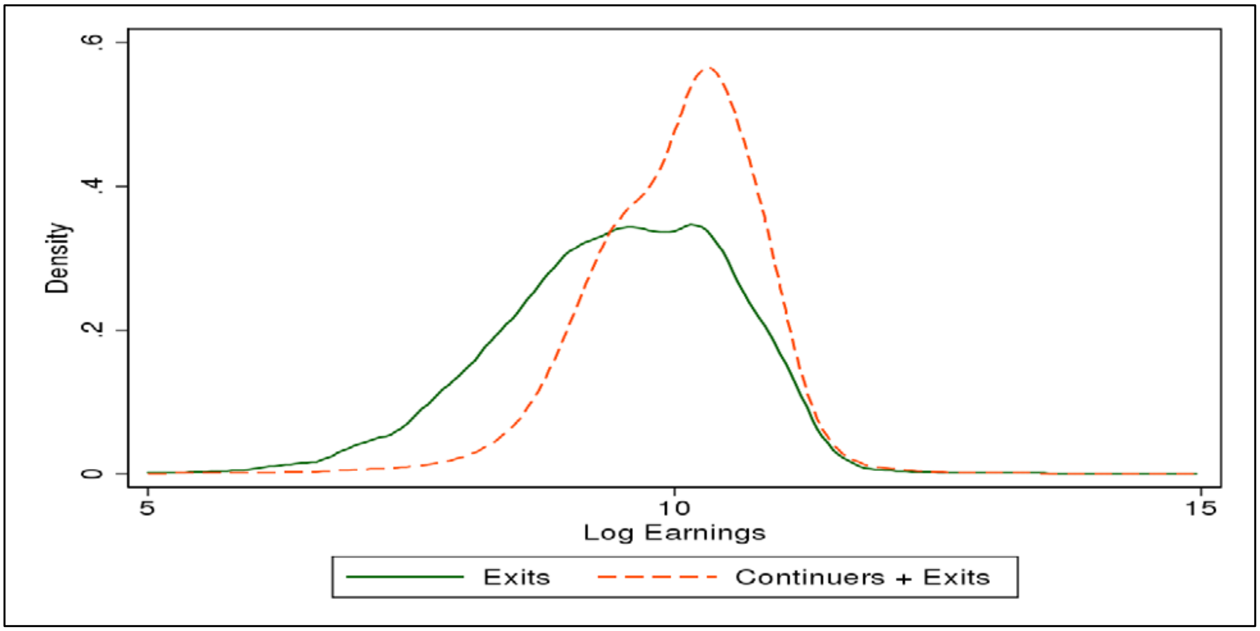
Note: The figure plots the kernel density estimates of the distribution of log average earnings for establishments that entered in 1992 or 1997 (solid line) versus the distribution for all establishments in operation in 1992 or 1997 (dashed line).

Figure 7 . Distribution of Earnings of Entering Establishments by Urban Density



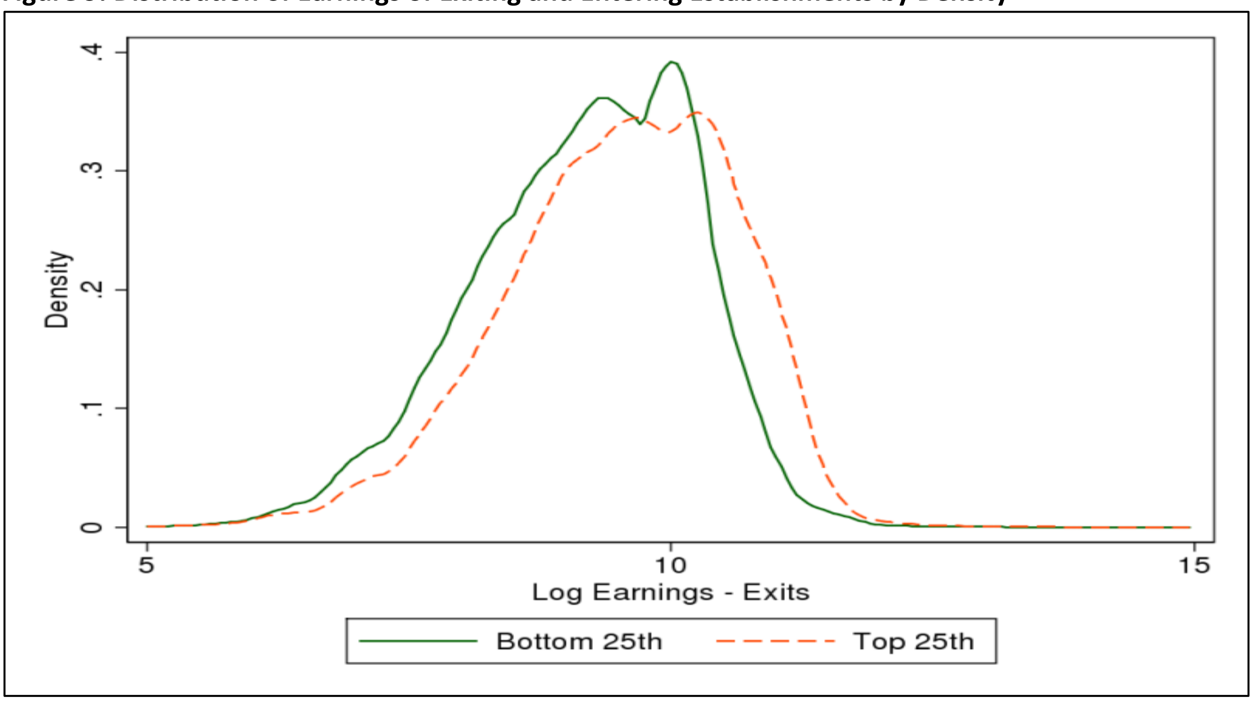
Note: The figure plots the kernel density estimates of the distribution of log average earnings for entering establishments in CBSAs in the bottom (solid line) and top (dashed line) quartiles of the density distribution, respectively. The top panel shows the unconditional distributions, while the bottom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

Figure 8. The Distribution of Earnings for Exiting and Surviving Establishments



Note: Figure plots the kernel density estimates of the distribution of log average earnings for establishments that exited in 1992 or 1997 (solid line) versus the distribution for all establishments in operation in 1991 and 1996 (dashed line).

Figure 9. Distribution of Earnings of Exiting and Entering Establishments by Density



Note: The figure plots the kernel density estimates of the distribution of log average earnings for exiting 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.