Creating Clusters of Entrepreneurship: Amenities, Education, and Historical Accidents

Edward L. Glaeser, William R. Kerr, Giacomo A. M. Ponzetto*
Harvard University
Cambridge, MA

April 2009

Abstract

Employment growth is strongly predicted by the number of establishments per worker, both across metropolitan areas and across industries within metropolitan areas, but there is little consensus on why this relationship exists. This paper presents a simple model which formalizes several different explanations for this phenomenon: lower fixed costs in some sectors in some places leads to both smaller firms and more entrepreneurship, higher profit margins leads to more competition and more entrepreneurship, a greater supply of entrepreneurial human capital leads to more firms and more growth and some places and sectors have evolved independent suppliers who abet competition and entrepreneurship. Evidence on returns does not supply the higher profit margins hypothesis, but all three other hypotheses receive some support in the data. Yet none of them can significantly explain the powerful correlation between firm size and later employment growth.

JEL Classification: J2, L0, L1, L2, L6, O3, R2.

Key Words: Entrepreneurship, Industrial Organization, Agglomeration, Cities.

*Comments are appreciated and can be sent to eglaeser@harvard.edu, wkerr@hbs.edu, and ponzetto@fas.harvard.edu. Kristina Tobio provided excellent research assistance. This research is supported by Harvard Business School, the Kauffman Foundation, the National Science Foundation, and the Innovation Policy and the Economy Group. The research in this paper was conducted while the authors were Special Sworn Status researchers of the US Census Bureau at the Boston Census Research Data Center (BRDC). Support for this research from NSF grant (ITR-0427889) is gratefully acknowledged. Research results and conclusions expressed are our own and do not necessarily reflect the views of the Census Bureau or NSF. This paper has been screened to insure that no confidential data are revealed.

1 Introduction

Economic growth is highly correlated with an abundance of entrepreneurial, small firms (Acs and Armington, 2006). As Figure 1 shows, when the number of firms per worker increases by 10% in 1977, employment growth between 1977 and 2000 increases by 9%. Within metropolitan areas, a 10% increase in firm size is associated with a 6% decline in employment growth over the same time period. The connection between small firms and area employment growth has been taken as evidence for the ability of competition to spur technological progress (Glaeser et al., 1992), the existence of a product cycle where growth is faster at earlier stages (Miracky, 1993), and the importance of entrepreneurship for area success (Glaeser, 2007). Any of these interpretations is compatible with the correlation between small firm size and employment growth, but the only thing that we can be sure of is that some areas seem like clusters of entrepreneurship, where firm sizes are small and growth is high, while others do not.

Silicon Valley is the archetypal example of such an entrepreneurship cluster. Saxenian's (1994) classic analysis of the region noted its large number of smaller, independent firms relative to Boston's Route 128 corridor. Following Chinitz (1961), Saxenian argued that these abundant small firms themselves caused the further outbreak of entrepreneurship by lowering the effective cost of entry with the development of independent suppliers, venture capitalists, entrepreneurial culture, and so on. This represents one interpretation of the connection between small firm size and later growth. A second interpretation, however, is that entrepreneurs, like Shockley, are drawn to Silicon Valley by its amenities and nice climate, which then leads to lots of little firms and a high rate of growth. Yet a third story would suggest that the returns to entrepreneurship were particularly high in the computer sector over the past 30 years.

This paper provides evidence about the relative importance of these hypotheses for explaining entrepreneurial clusters. We begin by documenting some basic facts about the connection between average establishment size and new employment growth through entrepreneurship using the Longitudinal Business Database. Section 2 confirms that there is a strong negative correlation between average initial establishment size and subsequent employment growth due to both new establishment formation and due to the expansion establishments of existing firms. As average firm size in 1992 increases by 10%, employment growth due to new startups over 1992-1999 declines by 7%. Employment growth due to facility expansion falls by almost 5%.

Moreover, the strongest correlations are within-region rather than across regions. There is a connection between average establishment size and employment growth at the city level, but there is a far stronger connection at the city-industry level. These patterns push towards theories that emphasize sector-region specific forces, rather than theories that emphasize region-wide variables that should have a common effect on all sectors. We also document that the reduction in new entry associated with larger firms, or a more concentrated industry in the region, comes primarily from a reduction in the employment growth in very small firms.[1]

Section 3 presents a framework, drawing on Chinitz (1961), that formalizes several possible explanations for these patterns. If the returns from starting a firm are higher in some areas, then there may be more start-ups and more small firms, at least in the short run, until the small firms have become bigger firms. In the model high returns come from more inelastic demand. If minimum scale is larger for firms in a sector, then that sector will have larger firms and fewer start-ups. The minimum required scale may be an exogenous attribute of the sector, or it might be an endogenous outcome depending on whether there are already independent suppliers for intermediate inputs. If amenities or historical accident draw an abundance of entrepreneurs in a sector to a particular area, then there will be more startups in that area and more firms per worker. .

Section 4 presents evidence on these hypotheses. To differentiate between the hypotheses that these areas have higher returns or lower costs, we use subsequent sales per worker among small firms as a proxy for the returns to entrepreneurship. We do not find that the strong relationship between initial industry structures and subsequent entry extends to these measures of entrepreneurial returns. This suggests that higher returns are unlikely to account for the observed link between lower initial establishment size and subsequent entry that we observe with our data. While some entrepreneurial clusters may be demand driven, the weight of our evidence suggests that supply side rationales regarding the cost of becoming an entrepreneur or are more important rationales.

Turning to the amenities or historical accidents rationales, it is first clear that the linkage must be deeper than simple industry-wide or city-wide forces. The powerful effects are robust to including both city and industry fixed effects, which rules out simple stories like entrepreneurs generally being attracted to urban areas with lots of amenities. Instead, as our model suggests, we look at interactions between area-level characteristics and industry-level characteristics. For example, the model suggests that entrepreneurship will be higher and firm size lower in high amenity places among industries with lower fixed costs. The evidence supports several hypotheses suggested by the model, but controlling for different forces does little to explain away the small firm size effect. Neither human capital characteristics of the area nor amenities can explain much of the firm size effect.

Our results document the remarkable correlation between average initial establishment size and subsequent employment growth due to startups. The evidence suggest that this correlation reflects a greater supply of entrepreneurship rather than greater returns to entrepreneurship in places with small firms. While evidence supports our model's explanations of the sorting of entrepreneurs and industries across space, the small establishment effects appears much stronger than these forces we have identified. We hope that future work will explore other mechanisms

[1] Prior empirical work on the Chinitz effect includes Drucker and Feser (2007), Rosenthal and Strange (2009), and Glaeser and Kerr (2009).

that can explain this important empirical regularity.

2 Clusters of Competition and Entrepreneurship

We begin with a description of the Longitudinal Business Database (LBD). We then document a set of stylized facts about employment growth due entrepreneurship. These descriptive pieces particularly focus on industry structure and labor intensity to guide and motivate our model development in Section 3.

2.1 LBD and US Entry Patterns

The LBD provides annual observations for every private-sector establishment with payroll from 1976 to 1999. Approximately four million establishments and 70 million employees are included each year. The Census Bureau data are an unparalleled laboratory for studying entrepreneurship rates and the life cycles of US firms. Sourced from US tax records and Census Bureau surveys, the micro-records document the universe of establishments and firms rather than a stratified random sample or published aggregate tabulations. In addition, the LBD lists physical locations of establishments rather than locations of incorporation, circumventing issues related to higher legal incorporations in states like Delaware.

The comprehensive nature of the LBD facilitates complete characterizations of entrepreneurial activity by cities and industries, types of firms, and establishment entry sizes. Each establishment is given a unique, time-invariant identifier that can be longitudinally tracked. This allows us to identify the year of entry for new startups or the opening of new plants by existing firms. We define entry as the first year that an establishment has positive employment. We only consider the first entry for cases where an establishment temporarily ceases operations (e.g., seasonal firms, major plant retoolings) and later re-enters the LBD. Second, the LBD assigns a firm identifier to each establishment that facilitates a linkage to other establishments in LBD. This firm hierarchy allows us to separate new startups from facility expansions by existing multi-unit firms.

Table 1 characterizes entry patterns in the manufacturing, services, retail trade, wholesale trade, mining, transportation, and construction sectors from 1977 to 1999. The first column refers to all new establishment formation. The second column looks only at those establishments that are not part of an existing firm in the database, which we define as entrepreneurship. The final column looks at new establishments which are part of an existing firm, which we frequently refer to as facility expansions.[2]

[2] Table 1 is drawn from Glaeser and Kerr (2009). Descriptive statistics for the current sample have not yet been disclosed. The current sample focuses on 1992-1999 to include greater coverage of services and financial sectors. Jarmin and Miranda (2002) describe the construction of the LBD. Sectors not included in the LBD are agriculture, forestry and fishing, public administration, and private households. We also exclude the US postal service, restaurants and food stores, hospitals, education services, and social services. These exclusions lower the relative share of services entrants.

Over the full sample period, there averaged 408,000 new establishments per annum with 3.8 million employments in the LBD. Single-unit startups account for 82% of new establishments but only 54% of new employment. Facility expansions are, on average, about four times larger than new startups. Figure 2 documents the distribution of establishment entry sizes for these two types. Over 75% of new startups begin with five or fewer employees, versus fewer than half for expansion establishments of existing firms. Less than 0.5% of independent startups begin with more than 100 workers, compared to 4% of expansion establishments.

Across industries, startups are concentrated in services (29%), retail trade (22%) and construction (20%). Facility expansions are concentrated in retail trade (42%), services (22%) and wholesale trade (17%). The growing region of the south has the most new establishment formations, and the regional patterns across the two classes of new establishments were quite similar. This uniformity, however, masks the agglomeration that frequently exists at the industry level. Well-known examples include the concentration of the automotive industry in Detroit, tobacco in Virginia and North Carolina, and high-tech entrepreneurship within regions like Silicon Valley and Boston's Route 128.[3]

2.2 Industry Structure and Entrepreneurship

Table 2 shows the basic fact that motivates this paper—the correlation between firm size and employment growth. We use both regions and metropolitan areas for spatial variation in this paper. While we prefer to analyze metropolitan areas, the city-level data become too thin for some of our variables when we use detailed industries. The dependent variable in the first three columns is the log employment growth in the region-industry due to new startups. The dependent variable for the second set of three columns is log employment growth in the region-industry due to new facility expansions that are part of existing firms.

Panel A uses the log of average firm size in the region-industry as the key independent variable. Panel B uses the Herfindahl-Hirschman Index (HHI) in the region-industry as our measure of industrial concentration. All regressions include the initial period's employment in the region as a control variable. For each industry, we exclude the region with lowest level of initial employment. This excluded region-industry is employed in the instrumental variable specifications. The cross of eight regions and 328 SIC3 industries yields 2,460 observations as not every region includes all industries. Estimations are unweighted and cluster standard errors by industry.

The first regression, in the upper left hand corner of the table, shows that the elasticity of employment growth in startups to initial employments is 0.96. This suggests that, holding mean establishment size constant, the number of startups scales almost one-for-one with existing employment. The elasticity of birth employment with respect to average establishment size in the

[3]Prior work on entry patterns using the Census Bureau data include Dunne et al. (1989a,b), Davis et al. (1996). Buenstorf and Klepper (2007) consider clusters and entry in the tire industry.

region-industry is -0.69. This relationship is both large and precisely estimated. It suggests that, holding initial employments constant, a 10% increase in average establishment size is associated with a 7% decline in the employment growth in new startups. These initial estimates control for region fixed effects (FEs) but not for industry FEs. Column 2 includes industry FEs so that all of the variation is coming from regional differences within an industry. The coefficient on average establishment size of -0.65 is remarkably close to that estimated in Column 1.

In the third regression, we instrument for observed average establishment size using the mean establishment size in the excluded region by industry. This instrument strategy only exploits industry-level variation, so we cannot include industry FEs. The estimated elasticities are again quite similar. These instrumental specifications suggest that the central relationship is not purely due to local feedback effects, where a high rate of growth in one particular region leads to an abundance of small firms in that place.

Panel B of Table 2 considers the log HHI index of concentration within each region-industry. While the model in the next section suggests using average establishment size to model industrial structure, there is also a long traditional of empirically modeling industrial structure through HHI metrics.[4] The results using this technique are quite similar to Panel A. A 10% increase in region-industry concentration in 1992 is associated with a 4% decline in employment due to new startups over 1992-1999. The coefficient on initial region-industry employment, however, is lower in this case. When not controlling for initial establishment size, there is a less than one-for-one relationship between initial employment and later growth through startups.

Column 2 of Panel B again models industry FEs. The coefficient is less stable than in the upper panel. The elasticity of startup employment to the HHI index continues to be negative and extremely significant, but it loses 50% of its economic magnitude compared to the first column. Column 3 instruments using the concentration level in the omitted region. The results here are quite similar to those in the first column.

Columns 4 to 6 of Table 2 consider employment growth from new facility expansions by multi-unit firms instead of new startups. These new establishments are not new entrepreneurship per se, but instead represent existing firms opening new production facilities, sales offices, and similar operations. Nevertheless, the new establishments represent more discontinuous events than simple employment growth at existing plants. Again, there is a strong negative effect of mean establishment size in the region-industry and subsequent employment growth due to facility expansions. The effect, however, is weaker than in the startup regressions. The results are basically unchanged when we include industry FEs or in the instrumental variables regression. These conclusions are also mirrored in Panel B's estimations using HHI concentration measures.

[4] The appendix also reports estimations using the share of employees in a region-industry working in establishments with 20 employees or fewer. This modelling strategy delivers similar results to mean establishment size or HHI concentration.

2.3 Entry Size Distribution

Table 3 quantifies how these effects differ across establishment entry sizes. Table 1 shows that most new establishments are quite small, while others have more than 100 workers. We separate out the employment growth due to new startups into groupings with 1-5, 6-20, 21-100, and 101+ workers in their first year of observation. Panel A again considers average firm size, while Panel B use the HHI concentration measure. These estimations only include region FEs, and the appendix reports similar patterns when industry FEs are also modelled.

A clear pattern exists across the entry size distribution. Larger average establishment size and greater industrial concentration both retard entrepreneurship the most among the smallest firms. For example, a 10% increase in mean establishment size is associated with a 12% reduction in new employment growth due to startups with five workers or fewer. The same increase in average firm size is associated, however, with less than a 1% reduction in new employment growth due to entering firms with more than 100 employees. The patterns across the columns show steady declines in elasticities as the size of new establishments increase. The impact for new firms with 6-20 workers is only slightly smaller than the impact for the smallest firms, while the elasticity for entrants with 21-100 employees is 50% smaller. Larger establishments and greater concentration are associated with a decrease in the number of smaller startups, but not a decrease in the number of larger startups.

2.4 City-Level Industrial Structures and Entrepreneurship

Tables 4A and 4B examine industrial structures in metropolitan areas. Table 4A focuses on the mean establishment size measure of industrial structure, while Table 4B considers the HHI concentration metrics. In both tables, the first two columns consider startup entry, while the last two columns consider facility expansions. Our data include 273 cities and 59 SIC2 industries. We move to the SIC2 level when considering city-industry combinations to reduce the number of zero-valued observations. Even at the higher industry level, however, not all city-industries are present which results in 14,471 observations. The other details of the regressions remain as before.

Column 1 only includes industry FEs so that we can consider the correlation between employment growth due to startups and metropolitan-level characteristics. This regression shows the large negative effect of greater average establishment size that is evident regionally is also evident by city-industry. A 10% increase in mean 1992 establishment size of a city-industry is associated an 8% reduction in employment growth due to startups over the ensuing eight years.

We also find a sizable and significant coefficient on the average establishment size in the metropolitan area as a whole. Holding a local sector's own establishment size constant, entrepreneurship increases when the surrounding city has greater numbers of small firms. The coefficients on initial employments are large and of similar magnitude to the mean establishment size, but with opposite signs. This pattern suggests that the growth of new startups is quite closely correlated to the number of existing establishments in the area. Column 2 includes city FEs. In this case, the coefficient on average establishment size falls to -0.67, and growth is not quite one-for-one with existing employment in the conditional estimation.

Columns 3 and 4 look at expansions of existing businesses. The impact of average establishment size in the city-industry remains robustly negative but is smaller than for startups. The role of initial employment in the city-industry is also quite similar to the first two columns. The biggest change is that the city-level mean establishment size no longer has a significant negative impact on employment growth.

2.5 Labor Intensity and Entrepreneurship

As a final descriptive piece, Table 5 documents that industries with high labor intensity have greater entrepreneurship rates. We believe that higher labor intensity is a reasonable proxy for lower overhead cost for starting a venture. This will be an element of the model described below. Panel B of Table 5 shows that controlling for this labor intensity does not substantially diminish the average establishment size effect. Thus, the earlier relationship between initial industry structure and subsequent entrepreneurship are not reflecting differing factor intensities.

In summary, these tables document the basic pattern that motivates this paper: there is an extremely strong connection between average establishment size and the entry of new establishments. This connection could be causal, but it seems just as likely to reflect omitted variables that turn some areas into clusters of entrepreneurship. The next section presents a model that explores different reasons why there could be a connection between average establishment size in an area and entry of new firms.

3 Theoretical Model

This model presents a formal treatment of entrepreneurship and industrial concentration. Our goal is to illustrate, within a single framework, a range of different explanations for the empirical implication and to produce added testable implications of these explanations. We first produce our most pared down model, which will enable us to look at the role of amenities, fixed costs and profitability in explaining entrepreneurship. We then extend the model so that it address multiple human capital levels, and then to allow for vertical integration.

We consider a closed economy with a perfectly inelastic factor supply. There are I cities, characterized by their exogenous endowments of real estate K_i, and by amenity levels a_i such that $a_i > a_{i+1}$ for all i. There is a continuum of industries $g \in [0, G]$, each of which produces a continuum of differentiated varieties.

Consumers have identical homothetic preferences defined over the amenities a of their city of residence, the amount of real estate K that they consume for housing, and their consumption

$q_g(\nu)$ of each variety in each industry. Specifically, we assume constant elasticity of substitution $\sigma(g) > 1$ across varieties in each sector, and an overall Cobb-Douglas utility function:

$$U = \log a + \rho \log K + (1-\rho) \int_0^G \beta(g) \left[\int_0^{n(g)} q_g(\nu)^{\frac{\sigma(g)-1}{\sigma(g)}} d\nu \right]^{\frac{\sigma(g)}{\sigma(g)-1}} dg, \tag{1}$$

with budget shares for consumption expenditures $\rho \in [0,1)$ and $\beta(g) > 0$ such that $\int_0^G \beta(g)\, dg = 1$.

Commodities are costlessly tradable across cities, while real estate is immobile. We assume for simplicity that it is owned by developers who reside in the same city where their property is located.[5] The economy comprises measure L of workers who are perfectly mobile across space: each supplies inelastically one unit of labor . Letting r_i denote the price of real estate and w_i the wage in city i, spatial equilibrium for workers requires that

$$\log a_i + \log w_i - \rho \log r_i = \log a_j + \log w_j - \rho \log r_j \text{ for all } i, j: \tag{2}$$

naturally, cities with higher amenities ($a_i > a_{i+1}$) have compensatingly lower wages ($w_i < w_{i+1}$) and higher rents ($r_i > r_{i+1}$).

Within each industry, omitting for the sake of brevity the index g, the preferences described by (1) imply the sectoral price index

$$P = \left[\int_0^n p(\nu)^{1-\sigma} d\nu \right]^{\frac{1}{1-\sigma}} \tag{3}$$

and the demand function for each variety

$$q(\nu) = RP^{\sigma-1} p(\nu)^{-\sigma}, \tag{4}$$

where $R = PQ$ is aggregate revenue in the sector, equal to a fraction $(1-\rho)\beta$ of total income.

Each variety ν is produced by a monopolistically competitive firm having increasing returns to scale. The cost function for a firm locating in city i is

$$c(q) = f r_i^{\kappa_f} w_i^{\lambda_f} + r_i^{\kappa_u} w_i^{\lambda_u} q. \tag{5}$$

This specification reflects a constant unit cost and an overhead requirement $f > 0$ that measures the extent of economies of scale. Technology is Cobb-Douglas, with cost shares different for real estate and for labor in each cost component, such that $\kappa_f + \lambda_f = 1$ for the overhead and $\kappa_u + \lambda_u = 1$ for the unit input requirement.

Monopolistic competition implies that each product has price

$$p = \frac{\sigma}{\sigma - 1} r_i^{\kappa_u} w_i^{\lambda_u}, \tag{6}$$

[5] Alternatively, we could allow for absentee ownership, and we would have rentiers living in extremely high-amenity cities where no production takes place.

and generically all firms in the sector prefer locating in the same city. Thus the sectoral price index is

$$P = n^{-\frac{1}{\sigma-1}} \frac{\sigma}{\sigma-1} r_i^{\kappa_u} w_i^{\lambda_u}, \tag{7}$$

and each firm earns profits

$$\pi = \frac{R}{\sigma n} - f r_i^{\kappa_f} w_i^{\lambda_f}. \tag{8}$$

Firms are created by entrepreneurs at a cost

$$C = F r_i^{\kappa_e} w_i^{\lambda_e}, \tag{9}$$

again resulting from Cobb-Douglas technology with cost shares $\kappa_e + \lambda_e = 1$. Relocating a firm is prohibitively costly, so its location is chosen once and for all at the moment of its creation; this rules out the presence of "nursery cities" à la Duranton and Puga (2001). Each firm has an exogenous probability δ of being forced out of the market in any period, e.g., because its product becomes obsolete and is no longer valued by consumers. In a steady state with a zero rate of time preference, free entry pins down the number of firms

$$n = \frac{R}{\sigma\left(\delta F r_i^{\kappa_e} w_i^{\lambda_e} + f r_i^{\kappa_f} w_i^{\lambda_f}\right)}. \tag{10}$$

Moreover, in each period and for each sector the aggregate profits of existing firms coincide with the aggregate payments to factors employed by entrepreneurs to create new varieties in the same industry.

Firm size measured by labor employment equals

$$\bar{L} = \left[\lambda_f + \lambda_u(\sigma-1)\right] f \left(\frac{r_i}{w_i}\right)^{\kappa_f} + \lambda_u(\sigma-1)\,\delta F \left(\frac{r_i}{w_i}\right)^{\kappa_e}, \tag{11}$$

and the fraction of workers in the sector employed by entrepreneurs creating new firms is

$$\iota = \left\{1 + \frac{\lambda_u}{\lambda_e}(\sigma-1) + \left[\frac{\lambda_f}{\lambda_e} + \frac{\lambda_u}{\lambda_e}(\sigma-1)\right]\frac{f}{\delta F}\left(\frac{r_i}{w_i}\right)^{\lambda_e-\lambda_f}\right\}^{-1}, \tag{12}$$

which together imply

$$\iota = \left[1 + \frac{\bar{L}}{\lambda_e \delta F}\left(\frac{w_i}{r_i}\right)^{\kappa_e}\right]^{-1}. \tag{13}$$

This condition delivers the amount of entrepreneurship in each industry. Inspection of the condition yields two potential explanations for this connection between the number of firms and the level of entrepreneurship. Relative employment in start-ups and the scale of existing firms move in opposite directions due to cross-sectoral variations in economies of scale and in product differentiation. We consider both in turn: proofs of the sorting of sectors into cities are provided in the appendix.

Proposition 1 *Suppose that industries differ by the degree of economies of scale $f(g)$.*

Within cities, sectors with higher economies of scale have fewer firms ($\partial n/\partial f < 0$), higher average employment per established firm ($\partial \bar{L}/\partial f > 0$), and a lower fraction of the work force employed by innovating entrepreneurs ($\partial \iota/\partial f < 0$).

Across cities, if and only if $\lambda_e > \lambda_f$ then sectors with higher economies of scale sort into cities with lower amenities, higher wages and a lower cost of real estate.

This proposition explains the small-firm effect on the basis of cost differences across sectors, focusing in particular on variations in the overhead input requirement. It is natural that sectors with higher economies of scale should have fewer, larger firms in equilibrium.

The first part of the proposition highlights that sectors with higher fixed costs will have a smaller number of firms and a lower level of entrepreneurship. As the costs of operating a firm rises, the equilibrium number of firms unsurprisingly declines, and there is also a decrease in the amount of entrepreneurial activity. As the costs of opening a new establishment rise, fewer people are interested in starting such establishments. The decrease in the steady-state number of firms entails a corresponding decline in the steady-state number of entrepreneurs, and thus of their employees. An increase in the overhead cost induces a less than proportional reduction in the equilibrium number of firms: thus the number of workers employed in managing existing firms increases.

This leads directly to the second part of the proposition. The spatial sorting of sectors is driven by relative factor intensities, since factor rewards move in opposite directions across cities. For sectors with higher economies of scale, the overall factor intensity is determined more by that of the overhead, and less by that of entrepreneurship. As a consequence, sectors with lower firm scale and greater innovation are attracted to high-amenity locations if and only if innovation is the more labor-intensive activity.

In addition or in alternative to supply-side differences, demand-side variation can also explain why entrepreneurship and small firms thrive in the same sectors.

Proposition 2 *Suppose that industries differ by the degree of product substitutability $\sigma(g)$, and therefore by the mark-up $\sigma(g)/[\sigma(g)-1]$.*

Within cities, sectors with higher product substitutability and lower mark-ups have fewer firms ($\partial n/\partial \sigma < 0$), higher average employment per established firm ($\partial \bar{L}/\partial \sigma > 0$), and a lower fraction of the work force employed by innovating entrepreneurs ($\partial \iota/\partial \sigma < 0$).

Across cities, if $\lambda_u < \min\{\lambda_e, \lambda_f\}$ then sectors with higher product substitutability and lower mark-ups sort into cities with lower amenities, higher wages and a lower cost of real estate.

The degree of product substitutability is the primary determinant of price markups and profit levels in this model. With CES preferences, the level of competitiveness in each sector is entirely determined by the degree of product differentiation. If varieties are highly substitutable, competition is intense and mark-ups are low.

The decline in mark-ups and profitability makes the sector less attractive to entrepreneurs, so fewer firms enter the market. At the same time, each firm must operate on high volumes and low margins to defray its fixed costs. Again, in a stationary equilibrium fewer firms mean fewer start-ups and lower employment in innovation. On the other hand, a decrease in mark-ups entails an inversely proportional increase in employment in direct production activities.

Hence lower product differentiation, which leads to lower profitability, implies that average factor intensity depends mostly on the unit input coefficient. If this is the least labor-intensive cost component, it follows that industries with harsher competitive conditions are particularly keen on inexpensive real estate, and thus locate in cities with lower amenities.

Thus we have shown how the connection between firms size and innovation can arise from exogenous variations in the underlying parameters that characterize supply and demand at the industry level. The same parameters can explain sorting of sectors into cities because of the fundamental difference between mobile labor and immovable real estate.

3.1 Heterogeneous Human Capital

The model can be extended to consider human capital as another determinant of entrepreneurship. Suppose that the economy is endowed with measure L of unskilled workers and H of skilled workers, and that technology is Cobb-Douglas in the two kinds of labor and real estate. Since both types of workers are perfectly mobile, the spatial equilibrium condition (2) implies that there is a single skill premium in the entire economy, and letting w_i denote the wage of unskilled workers in city i that of skilled workers is hw_i.

The cost function for a producer can be rewritten

$$c(q) = f r_i^{\kappa_f} w_i^{\lambda_f + \eta_f} h^{\eta_f} + r^{\kappa_u} w^{\lambda_u + \eta_u} h^{\eta_u} q, \tag{14}$$

with cost shares $\kappa_f + \lambda_f + \eta_f = 1$ and $\kappa_u + \lambda_u + \eta_u = 1$; while for an entrepreneur the entry cost becomes

$$C = F r_i^{\kappa_e} w_i^{\lambda_e + \eta_e} h^{\eta_e}, \tag{15}$$

with cost shares $\kappa_e + \lambda_e + \eta_e = 1$. In equilibrium, the ratio of skilled to unskilled workers in each industry is determined by the skill intensity of each cost component, and by their relative importance in the sector. In particular, we can establish the following result.

Proposition 3 *Suppose that skill intensity across cost components is ranked so that $\eta_e/\lambda_e > \eta_f/\lambda_f > \eta_u/\lambda_u$. Let industries differ either by the degree of economies of scale $f(g)$, or by the degree of product substitutability $\sigma(g)$. Within cities, relatively more skilled workers are then employed in sectors with more firms, lower average employment per established firm, and a higher fraction of the work force employed by innovating entrepreneurs ($\partial(H/L)/\partial f < 0$ and $\partial(H/L)/\partial\sigma < 0$).*

The proposition formalizes the intuition that entrepreneurship and entrepreneurs go together. The hypothesis is simply that innovation is the most skill-intensive activity, and that production workers are on average the least endowed with human capital. If this is the case, then naturally the same industry characteristics that increase entrepreneurship and reduce firm size also increase the overall skill intensity of the sector, and lead it to employ a higher share of skilled workers in equilibrium.

In this model, mobility is endogenous and workers are always in a spatial equilibrium. As such, there is no way for endowments of human capital to lead to more entrepreneurship. To address this type of exogenous sorting, we would need to drop the spatial equilibrium assumption and assume that workers were either fixed or tied to an area by historical accident.

3.2 Vertical Integration and the Chinitz Hypothesis

A more significant extension of the model involves going beyond exogenous determinants of firm size and entrepreneurship to explain their negative correlation by an endogenous channel. In the spirit of Chinitz (1961), we focus on the choice of firm organization. One of Chinitz's core ideas was that entrepreneurship would be higher in places that had abundant suppliers. In this model, we endogenize the decision to integrate suppliers and examine the implications of integration decisions on later entrepreneurship.

In this structure, firms specialize in one stage of the production process and engage in outsourcing if they enter a market that already hosts a number of upstream and downstream firms that could become partners for the new entrepreneur. Instead, if existing producers are vertically integrated, newcomers will perceive a need to enter as an equally integrated firm. These considerations in turn affect the equilibrium level of entrepreneurship, and we show that they also account for a link between a multitude of smaller firms and higher rates of entry.

Formally, we follow Grossman and Helpman's (2002) model of integration versus outsourcing in industry equilibrium. The production process of each differentiated variety requires two stages of production, which can be carried within an integrated firm or by outsourcing. The operation of integrated firms is described by the baseline model above, and yields profits

$$\pi_v = \frac{1}{\sigma}\left(\frac{\sigma-1}{\sigma}\frac{P}{r_i^{\kappa_u}w_i^{\lambda_u}}\right)^{\sigma-1}R - fr_i^{\kappa_f}w_i^{\lambda_f}.$$

Alternatively, firms may operate as specialized producers of final goods, with an overhead f_s; or as specialized producers of intermediates, with an overhead f_m. The relationship between the two is characterized by costly search and incomplete contracts.

After the overhead costs are incurred, each firm must search for a partner. The probability of finding one is described by a matching function that has constant returns to scale. If there are m and s specialized producers respectively of intermediates and of final goods in the market, the probability of a match is $\mu(s/m)$ for each specialized producer of intermediates, and $\mu(s/m)\,m/s$

for each specialized producer of final goods: the former is increasing and the latter decreasing in the ratio s/m.

Once a match takes place, the two partners fully specialize to each other's technology. The intermediate-good supplier produces a quantity $q(\nu)$ of the specialized intermediate, and his unit input requirement is α times that of an integrated firm. After all costs have been sunk, the two partners bargain. The final-good producer can turn each unit of the specialized intermediate into one unit of the final good. Otherwise, both parties have an outside option of zero. The bargaining share of the producer of intermediates is ω.

As a consequence, the price of each final good ν sold by specialized producers is

$$p_O = \frac{\sigma}{\sigma - 1} \frac{\alpha}{\omega} r^{\kappa_u} w^{1-\kappa_u}. \tag{16}$$

Expected profits for each specialized intermediate-good producer are

$$\pi_m = \mu\left(\frac{s}{m}\right) \frac{\omega}{\sigma} \left(\frac{\sigma - 1}{\sigma} \frac{\omega}{\alpha} \frac{P}{r_i^{\kappa_u} w_i^{\lambda_u}}\right)^{\sigma - 1} R - f_m r_i^{\kappa_f} w_i^{\lambda_f}, \tag{17}$$

and for each final-good producer

$$\pi_s = \mu\left(\frac{s}{m}\right) \frac{m}{s} (1 - \omega) \left(\frac{\sigma - 1}{\sigma} \frac{\omega}{\alpha} \frac{P}{r_i^{\kappa_u} w_i^{\lambda_u}}\right)^{\sigma - 1} R - f_s r_i^{\kappa_f} w_i^{\lambda_f}. \tag{18}$$

Firms of all types are hit by a fatal shock with the same constant hazard rate δ. The cost of entrepreneurship for a specialized producer of final goods is F_s and for a specialized producer of intermediates it is F_m. Free entry implies that

$$\begin{cases} \pi_m \leq \delta F_m r_i^{\kappa_e} w_i^{\lambda_e} \\ \pi_s \leq \delta F_s r_i^{\kappa_e} w_i^{\lambda_e} \end{cases}, \tag{19}$$

which must hold with equality for specialization to be an equilibrium organizational form; and similarly

$$\pi_v \leq \delta F r_i^{\kappa_e} w_i^{\lambda_e}, \tag{20}$$

which must hold with equality for integration to be an equilibrium organizational form. Generically, all firms in a sector prefer the same organizational form.

Since we are interested in the effect of organizational choice on entrepreneurship through channels other than the size of the overhead, which was already the focus of proposition 1, we assume that

$$\frac{f_m}{f} = \frac{F_m}{F} = \phi_m < 1 \text{ and } \frac{f_s}{f} = \frac{F_s}{F} = \phi_s < 1. \tag{21}$$

In any equilibrium with specialization

$$\frac{s}{m} = \sigma \frac{1 - \omega}{\omega} \frac{\phi_m}{\phi_s} \tag{22}$$

and an equilibrium in which all producers are specialized exists if and only if

$$\alpha^{\sigma-1}\phi_m \leq \omega^\sigma \mu\left(\sigma\frac{1-\omega}{\omega}\frac{\phi_m}{\phi_s}\right). \tag{23}$$

On the other hand, an equilibrium in which all producers are integrated always exists, since a single specialized producer could never find a partner to operate profitably; however, it is stable if and only if the outsourcing equilibrium does not exist.

In the outsourcing equilibrium there are

$$m = \frac{\omega R}{\sigma\phi_m\left(fr_i^{\kappa_f}w_i^{\lambda_f} + \delta F r_i^{\kappa_e}w_i^{\lambda_e}\right)} \tag{24}$$

specialized producers of intermediates and

$$s = \frac{(1-\omega)R}{\phi_s\left(r_i^{\kappa_f}w_i^{\lambda_f} + \delta F r_i^{\kappa_e}w_i^{\lambda_e}\right)} \tag{25}$$

specialized producers of final goods.

Firm size measured by labor employment equals

$$\bar{L}_O = \frac{\left\{\lambda_f\left[\omega+\sigma(1-\omega)\right] + \lambda_u(\sigma-1)\omega\right\} f\left(\frac{r_i}{w_i}\right)^{\kappa_f} + \lambda_u(\sigma-1)\omega\delta F\left(\frac{r_i}{w_i}\right)^{\kappa_e}}{\frac{\omega}{\phi_m}+\frac{\sigma(1-\omega)}{\phi_s}} < \bar{L} \tag{26}$$

and the fraction of workers in the sector employed by entrepreneurs creating new firms is

$$\iota_O = \left\{\left[1+\frac{(\sigma-1)\omega}{\omega+\sigma(1-\omega)}\frac{\lambda_u}{\lambda_e}\right] + \left[\frac{\lambda_f}{\lambda_e}+\frac{(\sigma-1)\omega}{\omega+\sigma(1-\omega)}\frac{\lambda_u}{\lambda_e}\right]\frac{f}{\delta F}\left(\frac{r_i}{w_i}\right)^{\lambda_e-\lambda_f}\right\}^{-1} > \iota. \tag{27}$$

We have therefore established the following result.

Proposition 4 *Compared to an equilibrium in which all firms are vertically integrated, an equilibrium in which all firms are specialized producers (if it exists) has higher mark-ups, lower average employment per established firm, and a higher fraction of the work force employed by innovating entrepreneurs.*

This proposition establishes that differences in the equilibrium organizational form across industries can account endogenously for the correlations that we previously explained exogenously.

The pervasive presence of specialized firms induces an increase in mark-ups as a direct consequence of incomplete contracting. Since production costs are incurred by a partner who will obtain only a fraction $\omega < 1$ of revenues in ex-post bargaining, output is proportionally lower and the mark-up is $1/\omega$ times the one charged by an integrated firm. The reduction in the average size of each firm is also intuitive: outsourcing tends to increase the number of firms both by separating stages of production and by reducing fixed costs for each firm.

Most important, outsourcing also yields an increase in entrepreneurship, spurred by the opportunity of matching with a complementary specialized producer. Many entrepreneurs are employed in creating firms to enter the matching market. However, not all are matched in equilibrium, and the output of those that are is reduced by contracting frictions. This implies that the share of workers allocated to entry, as opposed to actual production, is higher than under vertical integration.

The equilibrium mode of organizational form is independent of location, since it is not a function of factor rewards. In fact, equilibrium selection is mostly determined by parameters specific to the integration decision. Outsourcing is obviously more likely when it involves greater cost reductions (ϕ_m, ϕ_s and α are low). Its likelihood is first increasing and then decreasing in ω on $(0, 1)$, which is intuitive since both types of specialized producers must have incentives to enter the market in order for outsourcing to be sustainable.

The only parameter that affects both the properties of the baseline equilibrium with vertical integration and the likelihood of outsourcing is the elasticity of substitution σ. While its effect is not unambiguous, the following case is of particular interest.

Proposition 5 *If $\log(\alpha/\omega) > 1$ then outsourcing is more likely in sectors with a low elasticity of substitution (σ), for any matching function $\mu(.)$ having constant returns to scale.*

The feasibility condition (23) implies that the likelihood of outsourcing is decreasing in σ if and only if

$$\sigma \log \frac{\alpha}{\omega} > \frac{\partial \log \mu}{\partial \log (s/m)} \in (0, 1) . \tag{28}$$

This is because changes in σ have two distinct effects. On the one hand, greater substitutability reduces the number m of firms that enter as specialized producers of intermediates and increases the probability that each of them successfully finds a match: this unambiguously favours outsourcing. On the other hand, for $\alpha > \omega$ outsourcing is sufficiently inefficient in the final stage of production that vertically integrated firms would charge lower prices: this tends to make them prevail when competition is tougher. For a sufficiently high value of the ratio α/ω, the latter effects is certain to dominate.

In this case, the results of proposition 2 are reinforced by the endogenous channel of firm organization. In particular, cities with higher amenity levels attract industries that are more likely to have an equilibrium with pervasive outsourcing, which are precisely the sectors that tend to have more numerous and smaller firms and a higher rate of entrepreneurship.

4 Amenities and Clusters of Entrepreneurship

We turn now to empirical evidence on the origins of clusters of entrepreneurship. We first consider the explanation that there exist very high returns to entrepreneurship in certain regions. We then present empirical evidence on the model's predictions regarding amenities and clusters of entrepreneurship. Throughout these exercises, we are assessing in part whether these forces can explain the small establishment size effect documented in Section 2.

4.1 High Returns to Entrepreneurship

Our first exercise tests between two broad classes of theories about the underlying connection of small establishments and employment growth due to startups. One class of theories suggests that this correlation exists because the returns to entrepreneurship are higher in places with smaller firms and abundant employment growth. Those higher returns could reflect an exogenous parameter such as the elasticity of substitution in the model, or they could reflect the supply of more able entrepreneurs. Alternatively, the returns to entrepreneurship in places with abundant small firms may be equal to or weaker than those elsewhere. This scenario would suggest that the strong correlation between initial industry structure and subsequent employment growth due to startups descends more from a reduction in the costs of entrepreneurship rather than an increase in the returns to entrepreneurship.

Table 6 presents evidence on these hypotheses using region-industry data from the manufacturing sector. We focus on the manufacturing sector as we can calculate for it the 1997 dollar value of shipments per worker by region-industry separately for startup firms and multi-unit firms in the Census of Manufacturers. We use this shipments per employee as a proxy for profitability, subject to include industry FEs that control for industry-level production techniques, and therefore the returns to entrepreneurship.

Columns 1 and 2 model log shipments per worker among single-unit firms as the dependent variable, while the last two columns consider the similar measure among multi-unit establishments. Column 1 does not find a strong relationship between average establishment size in 1992 of the region-industry and the value per worker subsequently evident in 1997. This weak explanatory power is both in economic magnitudes and in statistical significance. There is some evidence of greater initial employment in areas with high subsequent, which could be evidence for an agglomeration effect or just that there is more employment in places where the returns to that employment are higher.

The limited evidence for abnormal subsequent shipments per worker also extends in Column 2 to the industry concentration measure. Likewise, the third and fourth columns find even weaker relationships when instead considering the labor returns among establishments of multi-unit firms. These patterns suggest that abnormal returns are not the driving force behind the observed relationships. Instead, the results point us to theories that emphasize either an abundance of entrepreneurial types in the area or a reduction in the costs of entrepreneurship.

4.2 Amenities and the Small Establishment Size Effect

The model featured only one city-level attribute—the level of amenities. Better amenities will drive up the price of land, attracting low fixed cost industries that tend to have higher levels of entrepreneurship. We start with these implications of the model and then ask whether amenity variables, or any others, can explain the strong connection between average establishment size and new establishment formation.

While there are certainly many man-made local amenities (e.g.,), we focus on predetermined climate amenities that can be taken as exogenous. We collect city-level data on coastal access, January temperature, July temperature, snowfall, and precipitation. While all of these variables can impact both production and consumption, they seem likely to primarily impact consumer well-being rather than the efficiency of firms.

We consolidate these variables into a single amenity index by using a housing price hedonic regression that is reported in the appendix. We regress the log average housing price in the metropolitan area in 1990 on these climate variables. Our primary specification uses just the log of each explanatory variable, and we have confirmed that we deliver very similar results using a piecewise linear function that is also reported in the appendix. The explanatory power of the two specifications are quite similar. The San Francisco Bay Area and Los Angeles are typically found to have the nicest consolidated amenities, while Little Rock, AR, and Tulsa, OK, are judged to have the weakest amenities.

A number of studies consider the productivity benefits that natural advantages can offer.[6] The appendix also documents how this amenities index is mostly uncorrelated or negatively correlated with productive natural advantages like the cost of electricity or coal, the availability of farmland, and the availability of timberland across states. This suggests that our constructed amenities index is unlikely to be reflecting production-related benefits to entrepreneurship.

Following the model, we look at the interaction between amenities and the degree of labor-intensity in the industry. Labor intensity is defined as the ratio between total payroll of the establishment and total shipments. In the model, this variable also captured the degree to which the industry was dependent on real estate, or other inputs, that become more expensive in high amenity places. As such, the model predicted that labor intensive industries would particularly locate in high amenity areas.

Table 7 provides our first look at the relationship between this amenity index and both employment and entrepreneurship. The first two regressions consider industrial specialization across cities. Column 1 regresses log total employment by city-industry on the city's amenity index, the industry's labor intensity, and their interaction. Variables are demeaned prior to interaction

[6] See Ellison and Glaeser (1999), Rosenthal and Strange (2001), Ellison et al. (2009), Glaeser and Kerr (2009), and Holmes and Lee (2009).

to restore main effects. There is a strong positive correlation between the amenity index and the overall level of employment in the metropolitan area. High amenity places generally attract people and firms. Labor intensive industries are also generally larger in size. The interaction of amenities and labor intensity is strongly positive, implying that more labor intensive industries are disproportionately located in high amenity cities. The model's predicted pattern of industrial specialization is thus generally supported and persists in Column 2's conditional estimations.

Columns 3 to 5 consider employment growth due to startup entry as the outcome measure. The coefficient on city-level amenities remains positive and quite significant, even controlling for initial employment in the city-industry. When we do not control for initial employment, the coefficient on the amenity index more than doubles in size. While the raw effect of amenities is quite positive, we do not find that the interaction works in the expected direction. There is a weak negative relationship between the interaction and employment growth due to startups. This negative effect flips sign when we do not control for initial employment in the sector. The interaction becomes smaller in magnitude and statistically insignificant in the fourth regression that includes city and industry FEs.

Column 5 further includes average establishment size as a control in the specification with industry and city fixed effects. The conditional framework does little to change the estimated interaction effect, but the interaction also does little to diminish the firm size effect, which is essentially identical to that estimated in Table 4. As a result, we conclude that the small establishment size effect extend beyond the sorting of labor intensive industries into high amenity areas.

Columns 6 and 7 examine the connection between amenities and employment growth due to entry of establishments that are part of existing firms. The raw effect of amenities in this regression is still positive, although it is small. The small coefficient on amenities reflects the fact that we are controlling for initial employment. Without that control, the amenity measure has a large coefficient of 0.8. In this case, the interaction is positive but statistically insignificant. In the seventh regression, we include area and industry fixed effects and the interaction again becomes positive. The last column includes log of average firm size in the city industry as a control. Again, the control does little to our estimated coefficients and the control remains similar in magnitude that estimated in Table 4. The firm size effect does not seem to be touched by the industrial sorting considered here.

Table 7 shows a robust relationship between the amenity index and both employment and employment growth. It is also true that the amenity index has a greater effect on employment levels in labor intensive industries. However, amenities do not have a significantly greater impact on growth for more labor intensive industries. There is a slight positive interaction effect for facility expansion and a slight negative effect for startups.

4.3 Education and the Small Establishment Size Effect

In the model, amenities were an exogenous force that shifted the supply of entrepreneurs across space. We now look at a less plausibly exogenous force, the share of the population with college degrees. If the skill share of the population in a city is reasonably fixed, and if skilled workers are particularly intensively used in entrepreneurship especially in skill intensive industries, then this supply of entrepreneurs could also explain the firm size effect.

In the first two regressions of Table 8, we examine again industry employment. We include both the amenity index and the share of the population with college degrees as control variables. The coefficient on the amenity index falls significantly when we control for the skill mix, perhaps reflecting the possibility that amenities increase employment in part by attracting more skilled people. We also find that there is a positive interaction between both variables and the skill intensiveness of the region, as measured by the share of the industry's workforce with college degrees at the national level. Unsurprisingly, industries that are skill intensive locate in skill heavy cities. Such industries also locate in amenity-heavy cities. The second regression includes area and industry FEs and continues to find this industrial specialization.

The third and fourth regressions use the employment growth due to independent establishment formation as the dependent variable. We control for initial employment and average firm size. Column 3 finds that places with more skills have more startup growth, especially in more skilled industries. Skills do seem to be a recipe for entrepreneurship, but this does meaningfully explain the average establishment size effect. The fourth column includes industry and city FEs. Again, the interaction between city skill shares and the skill share in the industry remain significant but there is no impact on the estimated firm size effect.

In last two regressions consider employment growth due to affiliated new establishments. In that case, we find a positive interaction between the skill intensiveness of the industry and both the area-level amenities and the area-level human capital. Skill intensive industries are expanding in places that are skill heavy and in places with good weather. However, once again these controls do little to eliminate the firm size effect.

5 Conclusion

The connection between small initial establishment size and employment growth is remarkably robust. This effect does not reflect industrial sorting on its own nor any purely city-level omitted variables. However, we remain unsure about whether this correlation represents some causal link between small establishment size and employment or whether this correlation reflects omitted variables which explain both outcomes.

The evidence on shipments per worker suggests that the returns to production are probably comparable, rather than higher, in city-industries with lots of small firms. This fact pushes us away from theories that emphasize the returns to entrepreneurship and towards theories that emphasize lower costs of entrepreneurship or greater supplies of entrepreneurship.

To look at the supply of entrepreneurship hypothesis, we consider at area level amenities and the share of the population with college degrees. Both of these variables are able to explain the location of employment in ways that fit our model. High skilled industries and labor intensive industries sort into high amenity areas. However, neither variable did much to eliminate the robust correlation between firm size and employment growth.

There are two natural interpretations of this connection between firm size and employment that remain. The effect actually may be causal, perhaps working through the supply of input provides as in Chinitz (1961). Alternatively, omitted variables my reduce the costs of entrepreneurship or raise the supply of entrepreneurs in cities with abundant small firms. While this work has been a preliminary foray into this topic, much future work is left to be done.

6 Appendix

6.1 Equilibrium Sorting

In equilibrium, free entry implies

$$\max_i \left\{ \left(r_i^{\kappa_u} w_i^{1-\kappa_u}\right)^{\sigma-1} \left(f r_i^{\kappa_f} w_i^{1-\kappa_f} + \delta F r_i^{\kappa_e} w_i^{1-\kappa_e}\right) \right\} = \left(\frac{\sigma-1}{\sigma} P\right)^{\sigma-1} \frac{R}{\sigma} \tag{29}$$

and generically firms strictly prefer a single location.

Consider two cities i and j such that $r_i < r_j$ and $w_i > w_j$, and an industry that is indifferent between the two. The indifference condition is

$$\left(r_i^{\kappa_u} w_i^{1-\kappa_u}\right)^{\sigma-1} \left(f r_i^{\kappa_f} w_i^{1-\kappa_f} + \delta F_e r_i^{\kappa_e} w_i^{1-\kappa_e}\right) = \left(r_j^{\kappa_u} w_j^{1-\kappa_u}\right)^{\sigma-1} \left(f r_j^{\kappa_f} w_j^{1-\kappa_f} + \delta F_e r_j^{\kappa_e} w_j^{1-\kappa_e}\right) : \tag{30}$$

if $\kappa_u > \max\{\kappa_e, \kappa_F\}$, this condition requires $r_i^{\kappa_u} w_i^{1-\kappa_u} < r_j^{\kappa_u} w_j^{1-\kappa_u}$, and thus city i is preferred by all sectors g with $\sigma(g) > \sigma$.

Indifference can also be written

$$\frac{f r_i^{\kappa_f-\kappa_e} + \delta F w_i^{\kappa_f-\kappa_e}}{f r_j^{\kappa_f-\kappa_e} + \delta F w_j^{\kappa_f-\kappa_e}} = \left(\frac{r_j}{r_i}\right)^{\kappa_e+\kappa_u(\sigma-1)} \left(\frac{w_j}{w_i}\right)^{1-\kappa_f+(1-\kappa_u)(\sigma-1)} : \tag{31}$$

if and only if $\kappa_f > \kappa_e$ the left-hand side is monotone decreasing in f, and thus city i is preferred by all sectors g with $f(g) > f$.

In either case, indexing sectors so that $f'(g) > 0$ or $\sigma'(g) > 0$, sectors $g \in (\bar{g}_{i-1}, \bar{g}_i]$ locate in city i, where $\bar{g}_0 = 0$ and $\bar{g}_I = G$, while the remaining $I-1$ thresholds $\bar{g}_i$ are endogenously determined. The spatial equilibrium condition for firms is

$$\frac{\left(r_i^{\kappa_u} w_i^{1-\kappa_u}\right)^{\sigma(\bar{g}_{i-1})-1} \left[f(\bar{g}_{i-1}) r_i^{\kappa_f} w_i^{1-\kappa_f} + \delta F_e r_i^{\kappa_e} w_i^{1-\kappa_e}\right]}{\left(r_{i-1}^{\kappa_u} w_{i-1}^{1-\kappa_u}\right)^{\sigma(\bar{g}_{i-1})-1} \left[f(\bar{g}_{i-1}) r_{i-1}^{\kappa_f} w_{i-1}^{1-\kappa_f} + \delta F_e r_{i-1}^{\kappa_e} w_{i-1}^{1-\kappa_e}\right]} = 1 \text{ for all } i > 1, \tag{32}$$

where it is understood that either $\sigma' > 0$ and $f' = 0$ or $\sigma = 0$ and $f' > 0$.

For each city i and each sector $g \in (\bar{g}_{i-1}, \bar{g}_i]$, aggregate factor payments satisfy

$$\begin{cases} r_i K(g) = \left\{ \frac{\kappa_e}{1+\frac{f(g)}{\delta F}\left(\frac{r_i}{w_i}\right)^{\kappa_f-\kappa_e}} + \kappa_f \frac{\frac{f(g)}{\delta F}\left(\frac{r_i}{w_i}\right)^{\kappa_f-\kappa_e}}{1+\frac{f(g)}{\delta F}\left(\frac{r_i}{w_i}\right)^{\kappa_f-\kappa_e}} + \kappa_v\left[\sigma(g)-1\right] \right\} \frac{(1-\rho)\beta(g)Y}{\sigma(g)} \\ w_i L(g) = \left\{ \frac{\lambda_e}{1+\frac{f(g)}{\delta F}\left(\frac{r_i}{w_i}\right)^{\kappa_f-\kappa_e}} + \lambda_f \frac{\frac{f(g)}{\delta F}\left(\frac{r_i}{w_i}\right)^{\kappa_f-\kappa_e}}{1+\frac{f(g)}{\delta F}\left(\frac{r_i}{w_i}\right)^{\kappa_f-\kappa_e}} + \lambda_v\left[\sigma(g)-1\right] \right\} \frac{(1-\rho)\beta(g)Y}{\sigma(g)} \end{cases}, \tag{33}$$

where aggregate income

$$Y = \sum_{i=1}^{I} \left[r_i K_i + w_i \int_{\bar{g}_{i-1}}^{\bar{g}_i} L(g)\, dg\right] \tag{34}$$

can be normalized to unity.

Considering that a fraction ρ of income in city i is spent on final consumption of real estate in the same city, the full-employment condition for real estate is

$$(1-\rho) r_i K_i = \int_{\bar{g}_{i-1}}^{\bar{g}_i} \left[r_i K(g) + \rho w_i L(g) \right] dg \text{ for all } i, \tag{35}$$

while that for labor is

$$L = \int_0^G L(g)\, dg. \tag{36}$$

Recalling the spatial equilibrium condition for workers (2), we have a system of $3I-1$ equations in as many unknowns: the $I-1$ cutoffs $\bar{g}_i$ and the I factor rewards (r_i, w_i):

$$\begin{cases} a_i w_i r_i^{-\rho} = a_{i-1} w_{i-1} r_{i-1}^{-\rho} \text{ for all } i>1 \\ \frac{r_i^{\kappa_e+[\sigma(\bar{g}_{i-1})-1]\kappa_u} w_i^{\lambda_e+[\sigma(\bar{g}_{i-1})-1]\lambda_u} \left[1+\frac{f(\bar{g}_{i-1})}{\delta F_e}\left(\frac{r_i}{w_i}\right)^{\kappa_f-\kappa_e}\right]}{r_{i-1}^{\kappa_e+[\sigma(\bar{g}_{i-1})-1]\kappa_u} w_{i-1}^{\lambda_e+[\sigma(\bar{g}_{i-1})-1]\lambda_u} \left[1+\frac{f(\bar{g}_{i-1})}{\delta F_e}\left(\frac{r_{i-1}}{w_{i-1}}\right)^{\kappa_f-\kappa_e}\right]} = 1 \text{ for all } i>1 \\ r_i K_i = \int_{\bar{g}_{i-1}}^{\bar{g}_i} \left\{ \frac{\kappa_e+\rho\lambda_e+\left(\kappa_f+\rho\lambda_f\right)\frac{f(g)}{\delta F}\left(\frac{r_i}{w_i}\right)^{\kappa_f-\kappa_e}}{1+\frac{f(g)}{\delta F}\left(\frac{r_i}{w_i}\right)^{\kappa_f-\kappa_e}} + (\kappa_v+\rho\lambda_v)\left[\sigma(g)-1\right] \right\} \frac{\beta(g)}{\sigma(g)} dg \text{ for all } i \\ L = \sum_{i=1}^I \frac{1-\rho}{w_i} \int_{\bar{g}_{i-1}}^{\bar{g}_i} \left\{ \frac{\lambda_e+\lambda_f\frac{f(g)}{\delta F}\left(\frac{r_i}{w_i}\right)^{\kappa_f-\kappa_e}}{1+\frac{f(g)}{\delta F}\left(\frac{r_i}{w_i}\right)^{\kappa_f-\kappa_e}} + \lambda_v\left[\sigma(g)-1\right] \right\} \frac{\beta(g)}{\sigma(g)} dg \end{cases} \tag{37}$$

It remains to be shown that this system in fact admits a unique solution.

6.2 Proof of Proposition 3

The ratio of skilled to unskilled workers in the industry equals

$$\frac{H}{L} = \frac{1}{h} \frac{\frac{\delta F r_i^{\kappa_e} w_i^{\lambda_e+\eta_e} h^{\eta_e}}{\delta F r_i^{\kappa_e} w_i^{\lambda_e+\eta_e} h^{\eta_e} + f r_i^{\kappa_f} w_i^{\lambda_f+\eta_f} h^{\eta_f}}\eta_e + \frac{f r_i^{\kappa_f} w_i^{\lambda_f+\eta_f} h^{\eta_f}}{\delta F r_i^{\kappa_e} w_i^{\lambda_e+\eta_e} h^{\eta_e} + f \Lambda^{\lambda_f} r^{\kappa_f} w^{1-\kappa_f}}\eta_f + (\sigma-1)\eta_u}{\frac{\delta F r_i^{\kappa_e} w_i^{\lambda_e+\eta_e} h^{\eta_e}}{\delta F r_i^{\kappa_e} w_i^{\lambda_e+\eta_e} h^{\eta_e} + f r_i^{\kappa_f} w_i^{\lambda_f+\eta_f} h^{\eta_f}}\lambda_e + \frac{f r_i^{\kappa_f} w_i^{\lambda_f+\eta_f} h^{\eta_f}}{\delta F r_i^{\kappa_e} w_i^{\lambda_e+\eta_e} h^{\eta_e} + f r_i^{\kappa_f} w_i^{\lambda_f+\eta_f} h^{\eta_f}}\lambda_f + (\sigma-1)\lambda_u}, \tag{38}$$

which is decreasing in f if and only if

$$\eta_e\lambda_f - \eta_f\lambda_e + \left[\left(\eta_e-\eta_f\right)\lambda_u - \eta_u\left(\lambda_e-\lambda_f\right)\right](\sigma-1) > 0, \tag{39}$$

and decreasing in σif and only if

$$\left(\frac{\eta_e}{\eta_u} - \frac{\lambda_e}{\lambda_u}\right)\delta F r_i^{\kappa_e} w_i^{\lambda_e+\eta_e} h^{\eta_e} + \left(\frac{\eta_f}{\eta_u} - \frac{\lambda_f}{\lambda_u}\right) f r_i^{\kappa_f} w_i^{\lambda_f+\eta_f} h^{\eta_f} > 0. \tag{40}$$

Both conditions are satisfied if, but not only if, $\eta_e/\lambda_e > \eta_f/\lambda_f > \eta_u/\lambda_u$.

References

[1] Acs, Zoltan, and Catherine Armington, *Entrepreneurship, Geography and American Economic Growth* (New York, NY: Cambridge University Press, 2006).

[2] Buenstorf, Guido, and Steven Klepper, "Why Does Entry Cluster Geographically? Evidence from the US Tire Industry", Working Paper (2007).

[3] Chinitz, Benjamin, "Contrasts in Agglomeration: New York and Pittsburgh", *American Economic Review* 51:2 (1961), 279-289.

[4] Davis, Steven, John Haltiwanger, and Scott Schuh, *Job Creation and Destruction* (Cambridge, MA: MIT Press, 1996).

[5] Delgado, Mercedes, Michael Porter, and Scott Stern, "Convergence, Clusters and Economic Performance", Working Paper (2008).

[6] Doms, Mark, Ethan Lewis, and Alicia Robb, "Local Labor Market Endowments, New Business Characteristics, and Performance", Working Paper (2008).

[7] Drucker, Joshua, and Edward Feser, "Regional Industrial Dominance, Agglomeration Economies, and Manufacturing Plant Productivity", CES Working Paper 07-31 (2007).

[8] Dunne, Timothy, Mark Roberts, and Larry Samuelson, "Patterns of Firm Entry and Exit in U.S. Manufacturing Industries", *RAND Journal of Economics* 19 (1989a), 495-515.

[9] Dunne, Timothy, Mark Roberts, and Larry Samuelson, "Plant Turnover and Gross Employment Flows in the U.S. Manufacturing Sector", *Journal of Labor Economics* 7 (1989b), 48-71.

[10] Duranton, Gilles, and Diego Puga, "Nursery Cities: Urban Diversity, Process Innovation, and the Life Cycle of Products", *American Economic Review* 91:5 (2001), 1454-1477.

[11] Ellison, Glenn, and Edward Glaeser, "The Geographic Concentration of Industry: Does Natural Advantage Explain Agglomeration?," *Papers and Proceedings, American Economic Review* 89 (1999), 311–316.

[12] Ellison, Glenn, Edward Glaeser, and William Kerr, "What Causes Industry Agglomeration? Evidence from Coagglomeration Patterns", *American Economic Review*, forthcoming (2009).

[13] Glaeser, Edward, "Entrepreneurship and the City", NBER Working Paper 13551 (2007).

[14] Glaeser, Edward, Hedi Kallal, Jose Scheinkman, and Andrei Shleifer, "Growth in Cities", *Journal of Political Economy* 100 (1992), 1126-1152.

[15] Glaeser, Edward, and William Kerr, "Local Industrial Conditions and Entrepreneurship: How Much of the Spatial Distribution Can We Explain?" *Journal of Economics and Management Strategy*, forthcoming (2009).

[16] Grossman, Gene, and Elhanan Helpman, "Integration versus Outsourcing in Industry Equilibrium", *Quarterly Journal of Economics* 117:1 (2002), 85-120.

[17] Holmes, Thomas, and Sanghoon Lee, "Economies of Density versus Natural Advantage: Crop Choice on the Back Forty", Working Paper (2009).

[18] Jacobs, Jane, *The Economy of Cities* (New York, NY: Vintage Books, 1970).

[19] Jarmin, Ron, and Javier Miranda, "The Longitudinal Business Database", CES Working Paper (2002).

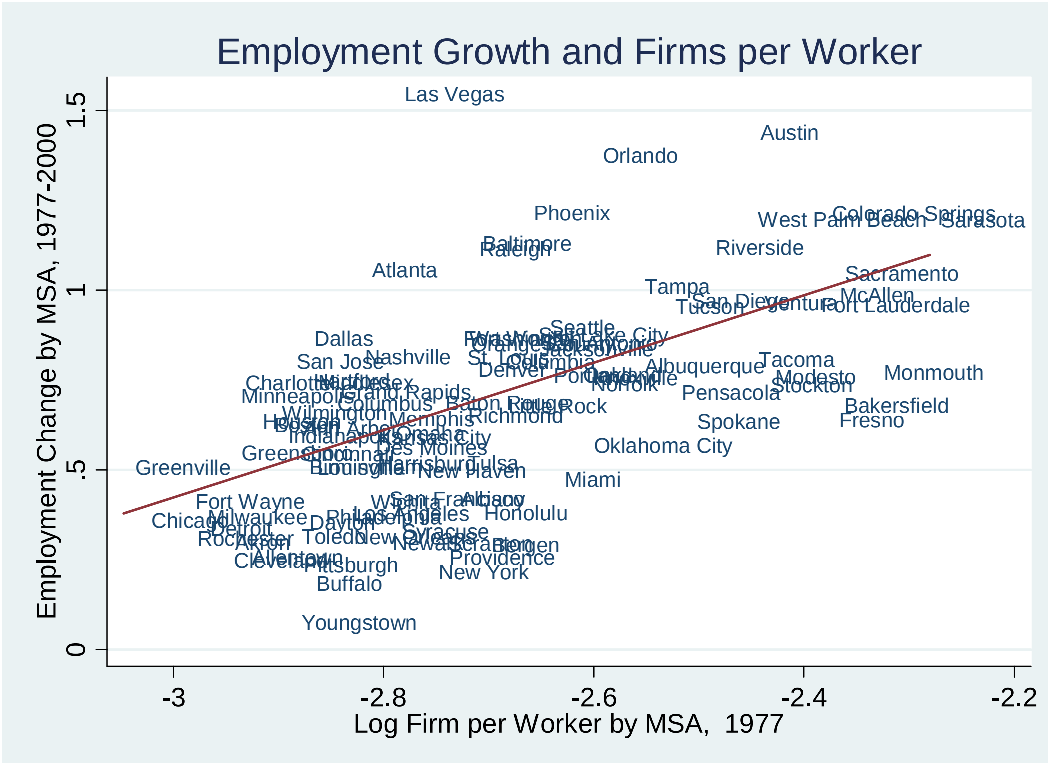
[20] Kerr, William, and Ramana Nanda, "Democratizing Entry: Banking Deregulation, Financing Constraints, and Entrepreneurship", *Journal of Financial Economics*, forthcoming (2009).

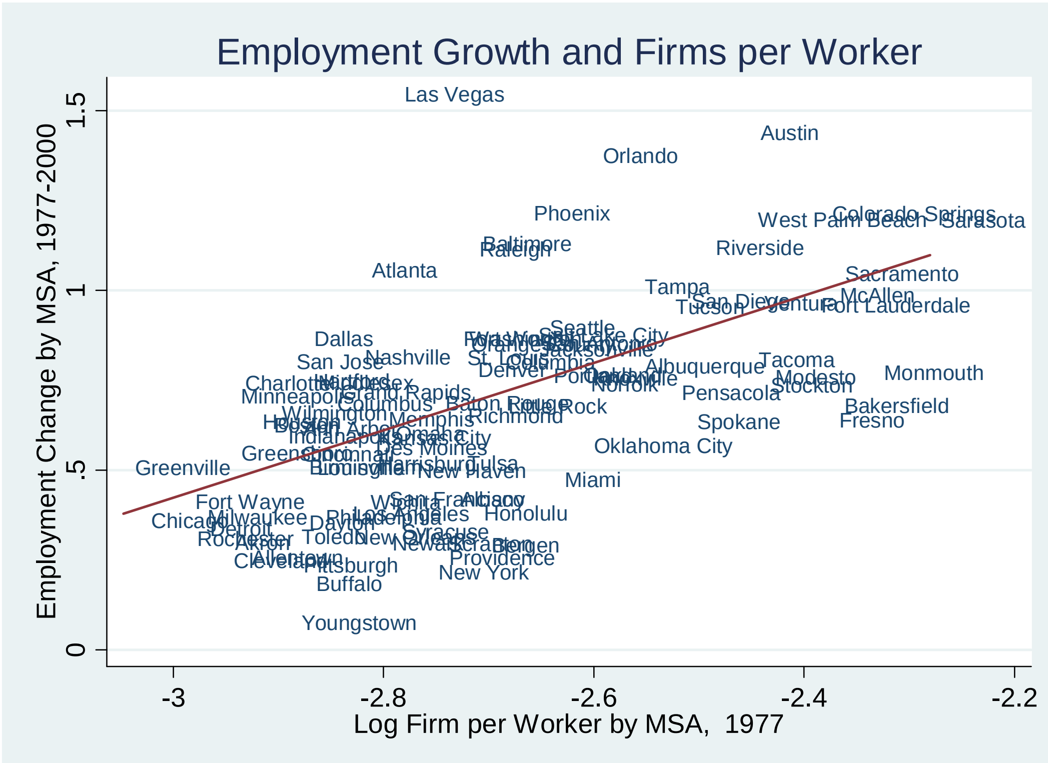
[21] Miracky, William, "Economic Growth and Business Cycles in Cities: The Role of Locational Externalities", MIT Thesis (1993).

[22] Rosenthal, Stuart, and William Strange, "Small Establishments/Big Effects: Agglomeration, Industrial Organization and Entrepreneurship", Working Paper (2009).

[23] Rosenthal, Stuart, and William Strange, "The Determinants of Agglomeration", *Journal of Urban Economics* 50 (2001), 191-229.

[24] Saxenian, AnnaLee, *Regional Advantage: Culture and Competition in Silicon Valley and Route 128* (Cambridge, MA: Harvard University Press, 1994).


Employment Growth and Firms per Worker
Employment Change by MSA, 1977-2000
Log Firm per Worker by MSA, 1977
0
.5
1
1.5
-3
-2.8
-2.6
-2.4
-2.2
Las Vegas
Austin
Orlando
Phoenix
Colorado Springs
West Palm Beach
Sarasota
Baltimore
Raleigh
Riverside
Atlanta
Sacramento
Tampa
McAllen
Fort Lauderdale
Tucson
Seattle
Dallas
Nashville
San Jose
Tacoma
Albuquerque
Monmouth
Denver
Modesto
Stockton
Norfolk
Pensacola
Bakersfield
Richmond
Fresno
Spokane
Memphis
Oklahoma City
Des Moines
Greenville
New Haven
Miami
Fort Wayne
Honolulu
Milwaukee
Chicago
Toledo
Syracuse
Providence
Pittsburgh
New York
Buffalo
Youngstown



Employment Growth and Firms per Worker
Employment Change by MSA, 1977-2000
Log Firm per Worker by MSA, 1977
0
.5
1
1.5
-3
-2.8
-2.6
-2.4
-2.2
Las Vegas
Austin
Orlando
Phoenix
Colorado Springs
West Palm Beach
Sarasota
Baltimore
Raleigh
Riverside
Atlanta
Sacramento
Tampa
McAllen
San Diego
Ventura
Tucson
Fort Lauderdale
Seattle
Dallas
Salt Lake City
Nashville
San Jose
St. Louis
Columbia
Jacksonville
Denver
Albuquerque
Tacoma
Monmouth
Modesto
Stockton
Norfolk
Pensacola
Charlotte
Minneapolis
Grand Rapids
Columbus
Baton Rouge
Little Rock
Bakersfield
Wilmington
Richmond
Memphis
Spokane
Fresno
Houston
Boston
Omaha
Kansas City
Indianapolis
Oklahoma City
Greensboro
Cincinnati
Des Moines
Greenville
Harrisburg
Tulsa
Louisville
New Haven
Miami
Fort Wayne
San Francisco
Albany
Milwaukee
Los Angeles
Honolulu
Chicago
Dayton
Philadelphia
Detroit
Syracuse
Rochester
Toledo
New Orleans
Akron
Newark
Bergen
Providence
Allentown
Cleveland
Pittsburgh
New York
Buffalo
Youngstown

Table 1: LBD Descriptive Statistics on US Entry Rates

	All Entering Establishments	Establishments of New Start-up Firms	Facility Expansions of Existing Firms
Mean Annual Entry Counts	407,783	335,807	71,976
Mean Annual Entry Empl.	3,811,409	2,081,801	1,729,608
Mean Annual Entry Size	9.3	6.2	24.0
Entry Counts by Entry Size			
1-5 Employees	70.3%	75.9%	44.4%
6-20 Employees	22.8%	19.8%	36.6%
21-100 Employees	5.8%	3.8%	14.9%
101+ Employees	1.1%	0.4%	4.1%
Entry Counts by Sector			
Manufacturing	9%	9%	6%
Services	28%	29%	22%
Wholesale Trade	12%	11%	17%
Retail Trade	25%	22%	42%
Mining	1%	1%	1%
Construction	17%	20%	1%
Transportation	7%	7%	10%
Entry Counts by Region			
Northeast	19%	20%	17%
South	36%	35%	37%
Midwest	22%	21%	24%
West Coast	24%	24%	22%

Notes: Descriptive statistics for entering establishments in the Longitudinal Business Database from 1977-1998. Jarmin and Miranda (2002) describe the construction of the LBD. Sectors not included in the LBD are agriculture, forestry and fishing, public administration, and private households. We also exclude the US postal service, financial services, restaurants and food stores, hospitals, education services, and social services. These exclusions lower the services share relative to other sectors. Incomplete LBD records require dropping 25 state-year files: 1978 (12 states), 1983 (4), 1984 (4), 1985 (1), 1986 (1), 1989 (1), and 1993 (2).

Table 2: Entry Rates and Regional Industrial Structure

	Start-Up Entry	Start-Up Entry	Start-Up Entry	Facility Expansions	Facility Expansions	Facility Expansions
	(1)	(2)	(3)	(4)	(5)	(6)
	Dep. Variable is Log Average Employment in Entering Establishments over 1992-1999 by Region-Industry					
	A. Entry and Average Establishment Size					
Log 1992 Total Employment in Region-Industry	0.963 (0.011)	0.812 (0.037)	0.960 (0.009)	1.101 (0.013)	0.952 (0.056)	1.093 (0.014)
Log 1992 Av. Establishment Size in Region-Industry	-0.689 (0.013)	-0.650 (0.046)	-0.654 (0.021)	-0.452 (0.031)	-0.355 (0.118)	-0.341 (0.032)
Adjusted R-Squared	0.82	0.93		0.72	0.90	
Region Fixed Effects	X	X	X	X	X	X
Industry Fixed Effects		X			X	
Estimation Technique	OLS	OLS	IV	OLS	OLS	IV
	B. Entry and HHI Concentration Index					
Log 1992 Total Employment in Region-Industry	0.651 (0.021)	0.465 (0.029)	0.528 (0.025)	0.911 (0.030)	0.754 (0.062)	0.829 (0.025)
Log 1992 HHI Concentration in Region-Industry	-0.403 (0.021)	-0.183 (0.022)	-0.592 (0.029)	-0.242 (0.030)	-0.125 (0.034)	-0.368 (0.025)
Adjusted R-Squared	0.75	0.93		0.69	0.89	
Region Fixed Effects	X	X	X	X	X	X
Industry Fixed Effects		X			X	
Estimation Technique	OLS	OLS	IV	OLS	OLS	IV

Notes: Estimations quantify the relationship between entry and industrial structure. The dependent variables are log entry employments of new firms or facility expansions by region-industry taken from the LBD. Entry employments are annual averages for region-industries over the 1992-1999 period. Regions are classified by the nine Census regions, and industries are classified at the SIC3 level (328 in total). The explanatory variables of total employments, average establishment size, and concentration are calculated from initial values in 1992 by region-industry. The region with the least industry employment is excluded for each industry in both OLS and IV specifications. IV regressions instrument for observed region-industry average establishment size or concentration with the 1992 level in the excluded region by industry. The first stage relationships are 0.899 (0.022) and 0.719 (0.014), respectively. Estimations report clustered standard errors, are unweighted, and have 2,460 observations. The decline in observations from the theoretical level of 2,624 is due to cases where an industry is not present in every region. Weighted regressions employing 1992 industry sizes as weights produce similar results.

Table 3: Entry Size Distribution and Regional Industrial Structure

	Total Entry	Entering Employment Of 1-5	6-20	21-100	101+
	(1)	(2)	(3)	(4)	(5)
	Dep. Variable is Log Average Employment in Entering Establishments over 1992-1999 by Region-Industry				
	A. Entry and Average Establishment Size				
Log 1992 Total Employment in Region-Industry	0.963	0.918	0.908	0.902	0.988
	(0.011)	(0.010)	(0.006)	(0.008)	(0.030)
Log 1992 Av. Establishment Size in Region-Industry	-0.689	-1.215	-0.953	-0.480	-0.043
	(0.013)	(0.012)	(0.011)	(0.013)	(0.025)
Adjusted R-Squared	0.82	0.86	0.85	0.74	0.49
Region Fixed Effects	X	X	X	X	X
	B. Entry and HHI Concentration Index				
Log 1992 Total Employment in Region-Industry	0.651	0.379	0.458	0.686	1.045
	(0.021)	(0.033)	(0.025)	(0.012)	(0.042)
Log 1992 HHI Concentration in Region-Industry	-0.403	-0.694	-0.585	-0.278	0.091
	(0.021)	(0.024)	(0.024)	(0.016)	(0.022)
Adjusted R-Squared	0.75	0.66	0.72	0.70	0.49
Region Fixed Effects	X	X	X	X	X

Notes: See Table 2. Estimations quantify the relationship between entry and industrial structure across the entry size distribution. Entering employments are for the first year of establishment observation.

Table 4A: Entry Rates and City-Level Industrial Structure

	Start-Up Entry	Start-Up Entry	Facility Expansions	Facility Expansions
	(1)	(2)	(3)	(4)
	Dep. Variable is Log Average Employment in Entering Establishments over 1992-1999 by City-Industry			
	Entry and Average Establishment Size			
Log 1992 Total Employment in City	0.237		0.340	
	(0.011)		(0.016)	
Log 1992 Total Employment in City-Industry	0.869	0.796	0.776	0.758
	(0.012)	(0.013)	(0.017)	(0.018)
Log 1992 Av. Establishment Size in City	-0.280		-0.014	
	(0.033)		(0.045)	
Log 1992 Av. Establishment Size in City-Industry	-0.807	-0.669	-0.529	-0.454
	(0.019)	(0.022)	(0.026)	(0.029)
Adjusted R-Squared	0.86	0.87	0.79	0.80
City Fixed Effects		X		X
Industry Fixed Effects	X	X	X	X

Notes: Estimations quantify the relationship between entry and city-level industrial structure. The dependent variables are log entry employments of new firms or facility expansions by city-industry taken from the LBD. Entry employments are annual averages for city-industries over the 1992-1999 period. Cities are classified by 273 PMSAs excluding AK and HI, and industries are classified at the SIC2 level (59 in total). The explanatory variables of total employments, average establishment size, and concentration are calculated from initial values in 1992 by city-industry. Estimations report robust standard errors, are unweighted, and have 14,471 observations. The decline in observations from the theoretical level of 16,107 is due to cases where an industry is in not present in every city. Weighted regressions employing an interaction of average industry size across cities with average size of industries within a city as weights produce similar results.

Table 4B: Entry Rates and City-Level Industrial Structure

	Start-Up Entry	Start-Up Entry	Facility Expansions	Facility Expansions
	(1)	(2)	(3)	(4)
	Dep. Variable is Log Average Employment in Entering Establishments over 1992-1999 by City-Industry			
	Entry and HHI Concentration Index			
Log 1992 Total Employment in City	0.391 (0.011)		0.458 (0.014)	
Log 1992 Total Employment in City-Industry	0.369 (0.009)	0.388 (0.009)	0.448 (0.012)	0.482 (0.011)
Log 1992 HHI Concentration in City	-0.100 (0.009)		-0.069 (0.012)	
Log 1992 HHI Concentration in City-Industry	-0.277 (0.012)	-0.227 (0.011)	-0.168 (0.015)	-0.142 (0.014)
Adjusted R-Squared	0.85	0.86	0.78	0.80
City Fixed Effects		X		X
Industry Fixed Effects	X	X	X	X

Notes: See Table 4A.

Table 5: Entry Rates and Industry Labor Intensity

	Start-Up Entry	Start-Up Entry	Facility Expansions	Facility Expansions
	(1)	(2)	(3)	(4)
	Dep. Variable is Log Average Employment in Entering Establishments over 1992-1999 by Region-Industry			
	A. Labor Intensity Only			
Log 1992 Total Employment in Region-Industry	0.915	0.917	1.067	1.067
	(0.014)	(0.013)	(0.015)	(0.014)
Log 1992 Labor Intensity in Region-Industry	0.396	0.562	-0.218	-0.204
	(0.038)	(0.033)	(0.042)	(0.031)
Adjusted R-Squared	0.69		0.68	
Region Fixed Effects	X	X	X	X
Estimation Technique	OLS	IV	OLS	IV
	B. Labor Intensity and Regional Industrial Structure			
Log 1992 Total Employment in Region-Industry	0.964	0.961	1.100	1.093
	(0.011)	(0.010)	(0.013)	(0.014)
Log 1992 Labor Intensity in Region-Industry	0.301	0.342	-0.284	-0.332
	(0.030)	(0.037)	(0.035)	(0.039)
Log 1992 Av. Establishment Size in Region-Industry	-0.674	-0.630	-0.465	-0.364
	(0.012)	(0.019)	(0.032)	(0.034)
Adjusted R-Squared	0.83		0.73	
Region Fixed Effects	X	X	X	X
Estimation Technique	OLS	IV	OLS	IV

Notes: See Table 2. Estimations quantify the relationship between entry and industry labor intensity. Labor intensity is measured as payroll divided by total sales. The region with the least industry employment is excluded for each industry in both OLS and IV specifications. IV regressions instrument for observed region-industry labor intensity with the 1992 intensity in the excluded region by industry. The first stage relationship is 0.804 (0.014).

Table 6: Manufacturing Labor Returns and Regional Industrial Structure

	Log 1997 Labor Returns in Single-Unit Firms in Region-Industry		Log 1997 Labor Returns in Multi-Unit Firms in Region-Industry	
	(1)	(2)	(3)	(4)
	Dep. Variable is Log 1997 Dollar Value of Shipments Over Employee Count by Region-Industry			
	Labor Returns and Regional Industry Structure			
Log 1992 Total Employment in Region-Industry	0.070 (0.027)	0.034 (0.022)	0.098 (0.029)	0.080 (0.027)
Log 1992 Av. Establishment Size in Region-Industry	-0.059 (0.040)		-0.026 (0.037)	
Log 1992 HHI Concentration in Region-Industry		-0.023 (0.017)		-0.015 (0.026)
Adjusted R-Squared	0.70	0.70	0.80	0.80
Region Fixed Effects	X	X	X	X
Industry Fixed Effects	X	X	X	X

Notes: See Table 2. Estimations quantify the relationship between industry returns to labor and industrial structure. The dependent variables are log dollar value of shipments in 1997 divided employee counts. Estimations consider 1059 observations from the manufacturing sector.

Table 7: Amenities, Industry Location, and Entrepreneurship

	Total Employment	Total Employment	Start-Up Entry	Start-Up Entry	Start-Up Entry	Facility Expansions	Facility Expansions	Facility Expansions
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Dependent Variable is Log Employment in Indicated Type of Establishment over 1992-1999 by City-Industry							
City-Level Amenities	0.965		0.498			0.177		
	(0.081)		(0.037)			(0.043)		
Industry Labor Intensity	0.590		0.528			-0.272		
	(0.038)		(0.017)			(0.018)		
City-Level Amenities x Industry Labor Intensity	0.345	0.345	-0.141	-0.037	-0.126	0.033	0.126	0.066
	(0.157)	(0.063)	(0.070)	(0.044)	(0.041)	(0.070)	(0.049)	(0.049)
Log 1992 Total Employment in City-Industry			0.760	0.416	0.790	0.832	0.497	0.747
			(0.005)	(0.009)	(0.013)	(0.006)	(0.011)	(0.018)
Log 1992 Av. Establishment Size in Region-Industry					-0.644			-0.443
					(0.022)			(0.029)
Adjusted R-Squared	0.03	0.79	0.63	0.86	0.87	0.59	0.79	0.79
City Fixed Effects		X		X	X		X	X
Industry Fixed Effects		X		X	X		X	X

Notes: Estimations quantify the relationships among city amenities, industry labor intensity, industry location, and entrepreneurship. The dependent variables in columns 1 and 2 are log employments by city-industry taken from the LBD. These specifications describe industrial location patterns. The dependent variables in columns 3-8 are log entry employments of new firms or facility expansions by city-industry. These specifications describe subsequent entrepreneurship rates. Entry employments are annual averages for city-industries over the 1992-1999 period. Average entry of less than one worker is recoded as one worker for these estimations. Cities are classified by 273 PMSAs excluding AK and HI, and industries are classified at the SIC2 level (59 in total). City-level amenities are calculated through 1990 housing prices and climate variables as described in the text and appendix. Industry labor intensity is measured as payroll divided by total sales. Total employments are calculated from initial values in 1992 by city-industry. Explanatory variables are demeaned prior to interaction to restore main effects. Estimations report robust standard errors and are unweighted. Columns 1 and 2 have 16,107 observations. Columns 3-8 have 14,471 observations after dropping city-industries where no initial employment existed. Weighted regressions employing an interaction of average industry size across cities with average size of industries within a city as weights produce similar results.

Table 8: Amenities, Education, Industry Location, and Entrepreneurship

	Total Employment	Total Employment	Start-Up Entry	Start-Up Entry	Facility Expansions	Facility Expansions
	(1)	(2)	(3)	(4)	(5)	(6)
	Dependent Variable is Log Employment in Indicated Type of Establishment over 1992-1999 by City-Industry					
City-Level Amenities	0.628		0.157		-0.093	
	(0.088)		(0.031)		(0.042)	
City-Level Bachelors' Share in 1990	0.180		0.299		0.341	
	(0.069)		(0.027)		(0.036)	
Industry Share of Workers with Bachelors' Education	0.512		-0.231		0.232	
	(0.034)		(0.012)		(0.020)	
City-Level Amenities x Industry Bach. Intensity	0.297	0.297	0.030	0.035	0.259	0.302
	(0.142)	(0.065)	(0.050)	(0.037)	(0.078)	(0.055)
City-Level Bach. Share x Industry Bach. Intensity	0.639	0.640	0.140	0.226	0.189	0.327
	(0.111)	(0.057)	(0.042)	(0.032)	(0.064)	(0.046)
Log 1992 Av. Establishment Size in Region-Industry			-0.860	-0.676	-0.514	-0.463
			(0.009)	(0.023)	(0.013)	(0.029)
Log 1992 Total Employment in City-Industry			0.957	0.784	0.956	0.743
			(0.004)	(0.014)	(0.006)	(0.019)
Adjusted R-Squared	0.04	0.79	0.78	0.87	0.67	0.78
City Fixed Effects		X		X		X
Industry Fixed Effects		X		X		X

Notes: See Table 7. City and industry education shares taken from 1990 Census.

App. Table: Amenities and Housing Prices

Dep. Variable is Log 1990 House Price by City	Linear Specification			Quintile Specification
Coastal Access	0.478	Coastal Access		0.507
	(0.063)			(0.069)
Log Average Annual	0.008	Snow Fall		-0.054
Snow Fall	(0.016)		Q2	(0.089)
Log Average Annual	-0.206			-0.163
Precipitation	(0.040)		Q3	(0.120)
Log Average January	0.174			-0.047
Temperature	(0.097)		Q4	(0.137)
Log Average July	-1.792			-0.027
Temperature	(0.335)		Q5 (Most)	(0.148)
		Precipitation		-0.092
			Q2	(0.071)
				-0.202
			Q3	(0.067)
				-0.132
			Q4	(0.067)
				-0.278
			Q5 (Most)	(0.067)
		January Temperature		0.099
			Q2	(0.074)
				0.290
			Q3	(0.100)
				0.267
			Q4	(0.128)
				0.225
			Q5 (Warmest)	(0.154)
		July Temperature		-0.031
			Q2	(0.067)
				-0.072
			Q3	(0.072)
				-0.275
			Q4	(0.081)
				-0.492
			Q5 (Warmest)	(0.088)
Adjusted R-Squared	0.40	Adjusted R-Squared		0.43

Notes: Estimations consider log housing prices by city taken from the 1990 Census. Predicted values from the regressions are used as composite amenities variables in main specifications. Estimations contain 275 observations and report robust standard errors.

App. Table: Climate-Based Amenities v. Traditional Natural Advantages

Pairwise Correlation of Climate Amenities and Other Natural Advantages by State		Pairwise Correlation of Labor Intensity and Other Dependencies by Industry	
Electricity Affordability	-0.301	Electricity Intensity	0.010
Natural Gas Affordability	-0.407	Natural Gas Intensity	-0.036
Coal Affordability	-0.476	Coal Intensity	-0.019
Farmland Percentage	-0.534	Livestock Intensity	-0.095
Timberland Percentage	0.222	Lumber Intensity	0.245
Population Density	0.340	Final Cons. Sales Intensity	-0.114

Notes: The first column presents pairwise correlations between calculated climate-based amenities and other forms of natural advantages at the state level excluding AK and HI. The second column presents pairwise correlations between labor intensity of industries, measured as payroll divided by sales, and other dependencies for manufacturing industries. These latter data are only available for the manufacturing sector.