



# Local Industrial Conditions and Entrepreneurship: How Much of the Spatial Distribution Can We Explain?

Edward L. Glaeser  
William R. Kerr

Working Paper

09-055

Copyright © 2008 by Edward L. Glaeser and William R. Kerr

Working papers are in draft form. This working paper is distributed for purposes of comment and discussion only. It may not be reproduced without permission of the copyright holder. Copies of working papers are available from the author.

# Local Industrial Conditions and Entrepreneurship: How Much of the Spatial Distribution Can We Explain?

Edward L. Glaeser  
Harvard University and NBER  
Cambridge MA

William R. Kerr\*  
Harvard Business School  
Boston MA

October 2008

## Abstract

Why are some places more entrepreneurial than others? We use Census Bureau data to study local determinants of manufacturing startups across cities and industries. Demographics have limited explanatory power. Overall levels of local customers and suppliers are only modestly important, but new entrants seem particularly drawn to areas with many smaller suppliers, as suggested by Chinitz (1961). Abundant workers in relevant occupations also strongly predict entry. These forces plus city and industry fixed effects explain between sixty and eighty percent of manufacturing entry. We use spatial distributions of natural cost advantages to address partially endogeneity concerns.

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

*Key Words:* Entrepreneurship, Industrial Organization, Agglomeration, Labor Markets, Input-Output Flows, Innovation, Research and Development, Patents.

---

\*Comments are appreciated and can be sent to [eglaeser@harvard.edu](mailto:eglaeser@harvard.edu) and [wkerr@hbs.edu](mailto:wkerr@hbs.edu). We are grateful to Jim Davis, Mercedes Delgado, Glenn Ellison, Thomas Hellmann, Ramana Nanda, Scott Stern, and two referees for advice on this research project. 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. This paper was first circulated in November 2007.

# 1 Introduction

Some places, like Silicon Valley, seem almost magically entrepreneurial with a new startup on every street corner. Other areas, like declining cities of the Rust Belt, appear equally starved of whatever local attributes make entrepreneurship more likely. This paper adds to the growing entrepreneurship literature by detailing local conditions that correlate with high entry rates of new manufacturing firms using the Census Bureau’s Longitudinal Business Database (LBD).

The LBD, which is described at length in Section 2, contains annual information on all US private-sector establishments between 1976 and 1999. We are able to distinguish plants that are part of larger firms versus those that stand alone. The new entry of stand-alone plants gives us our measure of entrepreneurship by city-industry. Section 2 paints a broad overview of firm entry patterns in the US economy, and manufacturing specifically, from 1977 onwards. We document extremely high levels of entry in US manufacturing as noted in previous work (e.g., Dunne et al. 1989a,b, Davis et al. 1996, Dumais et al. 2002).

Section 3 discusses our theories and explanatory variables for local entry conditions. Perhaps the simplest theory, emphasized by Glaeser (2007) among others, is that cities are more entrepreneurial if they have people whose demographics incline them towards entrepreneurship. We measure these demographics with age and education levels of cities. Another simple theory emphasizes innate cost advantages of particular regions for certain industries, such as coastal access for export industries or cheap electricity for aluminum production.

In addition to demographics and natural advantages, incumbent industrial structures of each city shape the availability and flow of goods, people, and ideas to new ventures. We test three theories that descend from agglomeration economies described by Marshall (1920). We look at whether entrepreneurship clusters around industries that are suppliers or customers, industries that employ similar types of labor, or industries that share ideas. These metrics are calculated at the city-industry level by uniting the distribution of incumbent firm types within each city with measures of the interdependencies among industries (e.g., Ellison et al. 2007). We also examine the Chinitz (1961) hypothesis more specifically. Chinitz argued that the presence of small, independent suppliers was particularly crucial for understanding why New York was so much more entrepreneurial than Pittsburgh. We test the Chinitz hypothesis by looking at whether new entry is more common when suppliers are smaller in size.

Section 4 comparatively assesses the explanatory power and importance of local conditions for entrepreneurship. We first test the ability of city-level characteristics to predict new manufacturing startups. We find limited evidence supporting the importance of demographics. In contrast to self-employment metrics (e.g., Glaeser 2007), manufacturing startups are not more

common in places with older or better-educated citizens. This makes sense given the scale and investment required, as well as the use of self employment by some older workers as a transition to retirement. Entry is, however, higher in cities with more workers between 20 and 40 years old. We also find little evidence for a “culture” of entrepreneurship. On the other hand, the Chinitz measure of small suppliers has very strong predictive power.

We then turn to our main regressions that include both city and industry fixed effects. The Chinitz measure is again the most important factor in these conditional estimations. Small suppliers predict new entrants, while general proximity to suppliers or customers is less important. We also find that the presence of industries that use the same type of labor is robustly important. This explanatory power for entrepreneurship holds when controlling for contemporaneous facility expansions by existing firms in the city-industry (e.g., Kerr and Nanda 2007). Looking across the entry size distribution, we find that Chinitz factors are most important for smaller entrants, while larger entrants more equally weight general input conditions. Technology and idea sharing also appear most important for smaller startups. Labor mix theories, on the other hand, receive equal emphasis throughout entrant size categories.

While the correlation between local industrial conditions and entry is impressive, it is certainly possible that firms and industries cluster in cities in anticipation of large amounts of entry. To address partially these endogeneity concerns, we turn to sixteen local characteristics that afford certain regions natural cost advantages for manufacturing industries (e.g., coastal access, timberland, energy prices, mean wages). Following Ellison and Glaeser (1999), predicted city-industry employment shares are developed by interacting these local cost advantages with factor intensities of industries in a non-linear least squares framework. This predicted spatial distribution of employment also predicts entrepreneurship well, highlighting the importance of basic cost considerations and natural advantages for explaining entry patterns. We then use these cost measures to predict Marshallian agglomeration economies by city-industry. We again find substantial evidence for the labor pooling and output markets rationales, while results for inputs and technology spillovers are not robust across specifications.

We conclude in Section 5. Many academics, policy makers, and business leaders stress the importance of local conditions for explaining spatial differences in entrepreneurship and economic development (e.g., Saxenian 1994, Acs and Armington 2006, Acs et al. 2008). This paper characterizes these entry relationships more precisely within the manufacturing sector. As manufacturing is a decreasing share of US employment, future research needs to explore other industrial sectors, too. Moreover, while our variables can explain between sixty and eighty percent of the spatial structure of manufacturing entrepreneurship, much of this comes from existing industry agglomeration and cost advantages. This suggests that there is still much more to learn about this mechanism for entrepreneurship in manufacturing as well. We hope that our empirical framework aids future inquiries in this vein.

## 2 Manufacturing Entrepreneurship

This section and the next outline the data and metrics employed in this study. We begin by discussing different techniques for measuring entrepreneurship. The strength and character of local entrepreneurship rates are then calculated from the LBD using new firm births. Drawing on Kerr and Nanda (2007), we describe the LBD’s structure and broad entry patterns that exist in multiple US industrial sectors. We then focus our attention on more detailed city-industry characteristics of manufacturing entrants that are considered in our empirical analyses. We discuss in Section 3 the metrics that we use to explain entrepreneurial patterns.

### 2.1 Measuring Entrepreneurship

Despite the extensive effort devoted to characterizing entrepreneurship, there is little consensus about the most appropriate metric. One approach associates entrepreneurship with the number of people leading independent enterprises. From this perspective, self-employment rates (e.g., Evans and Jovanovic 1989, Blanchflower and Oswald 1998) and average firm size (e.g., Glaeser 2007) are plausible measures. However, self employment weights small-scale, independent operators very heavily vis-a-vis high-growth entrepreneurship. This can be seen in self-employment rankings that list West Palm Beach, FL, as America’s most entrepreneurial city but place San Jose, CA, near the bottom.<sup>1</sup> Average firm size suffers less from this particular problem, but this metric captures little of the dynamic aspects of entrepreneurship and may reflect competition as much as entrepreneurship itself.

Understandably, many researchers are instead drawn to metrics that are more tightly connected to the dynamic nature of entrepreneurship. One approach focuses on startups within a single industry so that finer characterizations and case studies can be made (e.g., Saxenian 1994, Feldman 2003). An alternative looks at new product introductions (e.g., Audretsch and Feldman 1996), venture capital placement, or the founding of new firms (e.g., Kerr and Nanda 2007, Rosenthal and Strange 2007). These dynamic measures of entrepreneurship are less available than self-employment rates, but they do seem closer to the spirit of entrepreneurship these studies are seeking to capture. Our paper follows in this latter tradition.

We measure entrepreneurship as the formation of new manufacturing firms — quantifying both the count of new firms and the employment within them during their first year of operation. We focus on new establishments that are independent from existing firms. Our decision to

---

<sup>1</sup>This pattern is also evident in country rankings. For example, Southern European countries (e.g., Portugal, Greece) rank very high on European self-employment scales but tend to have very small venture capital markets. On the other hand, Scandinavian countries rank low on self-employment indices but have been among the most successful European countries in attracting venture capital investments (e.g., Bozkaya and Kerr 2007).

exclude establishments that are connected with existing firms is not without consequence. After all, entrepreneurial activity does take place within firms. Our sense, however, is that the entry of new firms is a better representation of entrepreneurship than facility expansions by existing manufacturing companies, which we will sometimes employ as a control. Data restrictions also limit us to firm births with payroll, which excludes hobby entrepreneurship, and we do not capture the share of output generated by new births.<sup>2</sup>

Our entry metric has a 0.36 and 0.66 correlation with the 2000 self-employment rates at the city and state levels, respectively. Correlation with average firm size is higher at -0.59 to -0.80. That is, smaller firm size is correlated with greater entry rates. Glaeser (2007) provides statistics on these alternative metrics and an analysis of their city-level determinants.

## 2.2 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. Manufacturing

---

<sup>2</sup>Our focus on manufacturing is also due to data constraints. While entry can be measured for other sectors by city, the Marshallian factors cannot be constructed. We hope that future research develops data and metrics appropriate for analyzing large-scale entry in other sectors.

accounts for just under 10% of the total entry; manufacturing, services, wholesale trade, and retail trade jointly account for 75%. Over 80% of the 400k new establishments opened in each year are new firm formations versus facility expansions.<sup>3</sup>

Figure 1a plots relative entry counts of both entrant types over time, with entry counts in 1977 to 1981 normalized to 100% for each group. While startups constitute the vast majority of new establishments, this time plot demonstrates that the relative increase in startup activity has consistently lagged that of expansion establishments since the early 1980s. There is only a 10% increase in the raw number of startup entrants over the twenty-year period, despite a 20% overall growth in LBD employment. Measured in terms of rates, Davis et al. (2006) document a substantial reduction in business entry and exit from the late 1970s to the late 1990s using the LBD. Figure 1b further displays the long-term sector decline for manufacturing. These metrics suggest that the US has become less entrepreneurial over the past three decades.

While startups account for the majority of new establishments, existing firms open new establishments at much larger sizes. New establishments of existing firms start on average with four times the employment of startups. Figure 2a documents the distribution of establishment entry sizes for these two types. Over three-quarters of new startups begin with five or fewer employees, versus fewer than half for expansion establishments of existing firms. The distribution differentials are even more pronounced in the capital-intensive manufacturing sector in Figure 2b.

The broad entry rates are fairly evenly spread across US regions, although this uniformity 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. Aggregate spatial distributions of startups versus facility expansions are relatively similar in Table 1.

Figure 3a depicts spatial differences in the fraction of manufacturing employment in entering firms. States are grouped by quintiles, and the darker shading indicates higher average entry shares. In line with strong population expansion and economic growth, western states and Florida are ranked at the upper end of the entry spectrum. Southern states are grouped around the middle, while manufacturing entrepreneurship rates are lowest among the Rust Belt states. These patterns also hold in Figure 3b's depiction of the fraction of manufacturing firms that are new entrants.

---

<sup>3</sup>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. Kerr and Nanda (2007) separately characterize entry in financial services. 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).

## 2.3 City-Industry Manufacturing Sample

Table 2 presents more detailed descriptive statistics for our manufacturing sample. We focus on 33,550 city-industry pairs that are formed by crossing 275 Primary Metropolitan Statistical Areas (PMSAs) with 122 SIC3 industries within manufacturing.<sup>4,5</sup> Rural areas are excluded, and we refer to PMSAs as cities in this paper for expositional ease. Table 2 presents the mean annual entry counts and entry employments of new firms by city-industry over the 1977-1999 period. Entry distributions are, of course, highly skewed. 32% of city-industry pairs do not have a single startup birth during the period, and about 60% of city-industry pairs have less than one entering employee on average per annum. The highest average startup entry employment is Electronic Components and Accessories (367) in San Jose, CA, while the highest entering firm count is Commercial Printing (275) in Los Angeles, CA.<sup>6</sup>

Table 2 also documents the industry employment share by city. These employments are used in Section 3 to estimate local industrial conditions. The most concentrated employment in our sample is Photographic Equipment and Supplies (386) in Rochester, NY, at over 50%. The excluded Fur Goods (237) industry is even more concentrated. We discuss below the endogeneity of using entry rates and industry employment calculated over the same 23-year horizon.

## 3 Determinants of Entrepreneurship

We now describe our measures of the determinants of entrepreneurship, beginning with the more exogenous determinants of demographics and natural advantages. We then turn to the agglomeration hypotheses of Marshall (1920) and Chinitz (1961). We conclude with entrepreneurial culture. Table 2 continues to document descriptive statistics.

---

<sup>4</sup>On the geography dimension, we map counties in the LBD to 317 PMSAs. We exclude 42 small PMSAs that are not separately identified in the 1990 Census of Population (required for explanatory variables). Results below are robust to instead considering Consolidated MSAs. CMSAs are subdivided into PMSAs for very large metropolitan areas (e.g., Chicago has six PMSAs within its CMSA). A PMSA is defined as a large urbanized county or a cluster of counties that demonstrate strong internal economic and social links in addition to close ties with the central core of the larger area.

<sup>5</sup>On the industry dimension, Tobacco (210s), Fur (237), and Search and Navigation Equipment (381) are excluded due to major industry reclassifications at the plant level that are difficult to interpret. The remainder of Apparel (230s), a portion of Printing and Publishing (277-279), and Secondary Non-Ferrous Metals (334) are also excluded due to poor data for constructing the Marshallian factors.

<sup>6</sup>Dunne et al. (1989a,b), McGuckin and Peck (1992), Davis et al. (1996), and Autor et al. (2007) provide additional details on entry patterns in the manufacturing sector. Dumais et al. (2002) and Ellison et al. (2007) provide further studies of the agglomeration and coagglomeration of startup and existing firm expansions, respectively.

### 3.1 Demographics

Demographics may influence the number of startups because certain types of people are more likely to be entrepreneurs. Higher self-employment rates, for example, are found in cities with older and better-educated populations. Demographics may also influence entry rates because certain types of workers are more likely to be desirable employees for young ventures. Both hypotheses would predict a link between demographics and entrepreneurial activity. Both hypotheses have been advanced by observers who argue that human capital policies that create and attract smart, entrepreneurial people are the key to local economic success. We develop from the 1990 Census of Populations simple statistics on the age distribution and education of the city’s workforce to test these factors.<sup>7</sup>

### 3.2 Natural Cost Advantages

A second hypothesis holds that some regions simply possess better natural environments for certain industries, and that entrepreneurship follows these natural cost advantages. Desert areas are inadequate hosts to the logging industry, coastal access is important for ship building or transporting very heavy products, and areas with cheap electricity attract aluminum producers. These cost advantages may lead to higher entry rates as well.

To model these advantages, we develop a predicted spatial distribution for each manufacturing industry based upon local cost advantages and industry traits. This work follows Ellison and Glaeser (1999), who model sixteen state-level characteristics that afford natural advantages in terms of natural resources, transportation costs, and labor inputs. Combining these cost differences with each industry’s intensity of factor use, Ellison and Glaeser (1999) estimate a spatial distribution of manufacturing activity that would be expected due to cost differences alone plus population distributions. They find that 20% of observed state-industry manufacturing activity can be explained through these mostly exogenous local factors.<sup>8</sup>

We extend this natural advantages estimation approach to the city-industry level, closely following the non-linear least squares approach of Ellison and Glaeser (1999). Where feasible and appropriate, we refine the earlier cost advantages to the city-level: coastal access, average

---

<sup>7</sup>The available evidence suggests that only a modest portion of entrepreneurs move to new cities in search of better local conditions. Michelacci and Silva (2007) find that Italian and US entrepreneurs are more likely to work in the region of their birth than employed workers. In a Danish sample collected by Dahl and Sorenson (2007), over 70% of new firms are founded in commuting regions where entrepreneurs were previously living. See also Figueiredo et al. (2002), Buenstorf and Klepper (2007), and Marx et al. (2007).

<sup>8</sup>Ellison and Glaeser (1999) suggest that this 20% share likely under-estimates the true portion of spatial agglomeration that can be explained through mostly fixed characteristics. Our explanatory power is higher than Ellison and Glaeser (1999) mostly due to our focus on three-digit rather than four-digit industries. Kim (1999) estimates natural regional advantages over a 100-year period.

manufacturing wages, education levels, population densities, and income shares. Other characteristics remain at the state level (e.g., farmland, timberland, various energy prices). The maximum predicted city-industry shares, defined as  $Nat\%_{ci}$ , are for consumer-focused industries in New York City at just over 10%. The weakest predicted share is for Blast Furnace and Basic Steel Products (331) into Anchorage, AK. The partial correlation of actual and predicted city-industry distributions is 0.42, a pseudo first stage between natural cost advantages and actual agglomeration.

### 3.3 Agglomeration Theories

New York City’s port and Pittsburgh’s coal mines would have attracted entrepreneurs regardless of the other firms located in those cities. In many cases, however, entrepreneurs are drawn by the existing industrial structure of a city. Entrepreneurs may cluster near potential customers or potential suppliers. New startups may draw ideas from neighboring firms. The composition and availability of early hires may be constrained by local labor pools. We now turn to these agglomeration economies of Marshall (1920).<sup>9</sup>

#### 3.3.1 Customer and Supplier Strength

The simplest agglomeration economy is that proximity to customers and suppliers reduces transportation costs and thereby increases productivity. The savings benefit of reduced shipping costs to distant consumers is the core agglomerative force of the new economic geography (e.g., Fujita et al. 1999). Where customers and suppliers are geographically separate, firms must trade-off distances. When production involves a large reduction in weight, it makes sense to produce close to raw materials similar to Chicago’s stockyards. When transporting finished products is quite difficult, it makes sense to produce downstream near the point of consumption. The difficulties of transporting refined sugar from the tropics in the nineteenth century led to sugar refinement in New York City, far from source plantations. In addition to shipment costs, Porter (1990) emphasizes that proximity to customers and suppliers can enhance innovation by increasing knowledge flows about which products are working and what new products are desired.

To test the importance of this mechanism, we measure the extent to which cities are rife with potential customers and suppliers for a new entrepreneur. We begin with the 1987 Benchmark Input-Output Accounts published by the Bureau of Economic Analysis (BEA). We define  $Input_{i \leftarrow k}$  as the share of industry  $i$ ’s inputs that come from industry  $k$ , and  $Output_{i \rightarrow k}$

---

<sup>9</sup>Several papers assess the relative importance of these determinants for industrial agglomeration, including Audretsch and Feldman (1996), Ellison and Glaeser (1997), Rosenthal and Strange (2001), Henderson (2003), and Ellison et al. (2007). These assessments are harder empirically than our current exercise due to the endogeneity of linkages that form between clustered firms. We take these linkages to be exogenous to startups in this paper.

as the share of industry  $i$ 's outputs that go to industry  $k$ . These measures run from zero (no input or output purchasing relationship exists) to one (full dependency on the paired industry). These shares are calculated relative to all input-output flows, including those to non-manufacturing industries or to final consumers. Customer and supplier flows are not symmetrical ( $Input_{i \leftarrow k} \neq Input_{k \rightarrow i}$ ). Moreover, differences in industry size and the importance of flows to or from non-manufacturing industries and final consumers result in asymmetries between pairwise customer and supplier dependencies ( $Input_{i \leftarrow k} \neq Output_{k \rightarrow i}$ ).<sup>10</sup>

Ellison et al. (2007) document large asymmetries that exist in these material flows. Approximately 70% of pairwise industrial combinations have an input-output dependency less than 0.01%. The strongest relative customer dependency is Leather Tanning and Finishing's (311) purchases from Meat Products (201) at 0.39 (i.e., 39% of 311's inputs come from 201). The highest absolute customer dependency (with a relative share of 23%) is Misc. Plastics Products' (308) purchases from Plastic Materials and Synthetics (282). The strongest relative output or supplier dependency is Public Building and Related Furniture's sales to Motor Vehicles and Equipment (371) at 82%. The highest absolute supplier dependency (with a relative share of 32%) is Plastic Materials and Synthetics' (282) sales to Misc. Plastics Products (308).

To summarize local conditions across industries within a city, we aggregate across all potential input suppliers. For a focal industry  $i$  in city  $c$ , we define the raw input opportunity as

$$Input_{ci} = - \sum_{k=1, \dots, I} \left| Input_{i \leftarrow k} - \frac{E_{kc}}{E_c} \right|, \quad (1)$$

where  $I$  indexes industries. This measure simply aggregates absolute deviations between the proportions of industrial inputs required by industry  $i$  and city  $c$ 's actual industrial composition during the sample period. The measure is orthogonal to city size, which we separately consider, and a negative value is taken so that the metric ranges between negative two (i.e., no inputs available in the local market) and zero (i.e., all inputs are available in the local market in precise proportions). The construction of  $Input_{ci}$  assumes that firms have limited ability to substitute across material inputs in their production processes.

To capture the relative strength of output relationships, we also define

$$Output_{ci} = \left[ \sum_{k=1, \dots, I} Output_{i \rightarrow k} \frac{E_{kc}}{E_k} \right] \cdot \left[ \sum_{k=1, \dots, I} Output_{\rightarrow k} \frac{E_{kc}}{E_k} \right]^{-1}. \quad (2)$$

The first bracketed term multiplies the national share of industry  $i$ 's output sales that go to industry  $k$  with the fraction of industry  $k$ 's employment in city  $c$ . By summing across industries,

---

<sup>10</sup>The "Use of Commodities by Industries" table provides commodity-level make and use flows for very detailed industries at the national level, which we aggregate to the SIC3 framework. While some commodities can partly be produced by other industries than the one associated with these commodities, we ignore this distinction and interpret the estimates as measuring how much of an industry's production is used as an input to other industries.

we measure the concentration of industrial sales opportunities for industry  $i$  in the focal market  $c$ . To maintain independence of market size, we normalize this measure through the second bracketed term in (2) that measures total potential industrial sales into the focal city. This measure takes on values between zero and one, with higher values suggesting greater selling opportunities. Unlike our input measure,  $Output_{ci}$  pools across industries that normally purchase goods from industry  $i$ . By measuring the aggregate strength of industrial sales opportunities in city  $c$ , the metric assumes that selling to one large industrial market is the same as selling smaller amounts to multiple industries.

Table 2 documents descriptive statistics for  $Input_{ci}$  and  $Output_{ci}$ . The best measured local supplier environments are for Hats, Caps, and Millinery’s (235) entry into the Carolinas and Communications Equipment’s (366) entry into San Jose, CA. The weakest input settings are associated with Petroleum Refining (291), Pulp Mills (261), and Carpets and Rugs (227) in multiple locales. The best measured local sales environments are for Motor Vehicles and Equipment’s (371) entry into Detroit, MI; a number of city-industry pairs are judged to offer poor industrial sales opportunities for firms.

$Input_{ci}$  and  $Output_{ci}$  condense large and diverse industrial structures for cities into manageable statistics of local industrial conditions. The metrics do have limitations, though. First, we do not capture potential customer or supplier interactions that exist beyond the local city, but perhaps within the state for example. Second, the metrics do not consider final consumers. In unconditional estimates, we separately model city populations. The metrics also suffer from endogeneity. There will surely be more input suppliers for industry  $i$  in a given city if that city attracts a steady flow of new entrants willing to buy the input suppliers’ products.

Chinitz (1961) also emphasized the role of input suppliers in his account of entrepreneurial differences between New York City and Pittsburgh. New ventures have many needs that must be met by the local economy, in contrast with larger incumbents that may source internally or at greater distances. Chinitz particularly stressed the interactions of startups with small, independent suppliers. Greater competition and smaller sizes for suppliers help new ventures source specialized inputs and avoid hold-up problems. Chinitz argued that the large, integrated steel firms of Pittsburgh depressed external supplier development; moreover, existing suppliers had limited interest in providing inputs to small businesses. By contrast, New York City’s much smaller firms, organized around the more decentralized garment industry that then dominated the city, were better suppliers to new firms. Jacobs (1970) also argued this perspective.

We test the Chinitz hypothesis — as distinct from the high-quality, general-input conditions of Marshall (1920) captured in (1) — through a metric that essentially calculates the average

firm size in city  $c$  in industries that typically supply a given industry  $i$ ,

$$Chinitz_{ci} = \sum_{k=1, \dots, I} \frac{Firms_{kc}}{E_{kc}} \frac{E_{kc}}{E_c} Input_{i \leftarrow k} = \sum_{k=1, \dots, I} \frac{Firms_{kc}}{E_c} Input_{i \leftarrow k}, \quad (3)$$

where  $Firms_{kc}$  is the count of firms. Higher values of this index indicate greater numbers of firms are typically providing input needs of new entrants, weighted by the importance of the inputs in question. The largest Chinitz measures are associated with Sawmills and Planing Mills (242).

### 3.3.2 Labor Market Strength

Labor may be the most important input into any new firm, and entrepreneurship is quite likely to be driven by the availability of a suitable labor force (e.g., Combes and Duranton 2006, Dahl and Klepper 2007). To some extent, basic demographics of an area (e.g., mean years of schooling) are informative about the suitability of the local labor force. But, these aggregate traits can also be quite blunt. Many industries require specialized occupations, and the share of college graduates in a city tells us little about the presence of such specialized workers. Zucker et al. (1998) describe the exceptional embodiment of human capital in specialized workers in the emergence of the US biotech industry.

The agglomeration of specialized workers and firms can occur through several channels. Marshall (1920) described how an agglomeration of workers and firms shields workers from firm-specific shocks. Workers can be more productive and better insured by moving from firms that are hit with negative shocks to better opportunities (e.g., Diamond and Simon 1990, Krugman 1991, Overman and Puga 2007). Larger labor pools further promote more efficient matches (e.g., Helsley and Strange 1990), and multiple firms protect workers against ex post appropriation of investments in human capital (e.g., Rotemberg and Saloner 2000). All of these mechanisms suggest that firms that employ similar types of workers will tend to locate near one another and that startups will benefit from thick local markets for their specific labor needs, either through heightened availability or lower wages.

We quantify the suitability of local labor markets by city-industry through an interaction of the incumbent manufacturing industrial structure of each city with the occupational labor requirements of each industry. Our data come from the 1987 National Industry-Occupation Employment Matrix (NIOEM) published by the Bureau of Labor Statistics (BLS). The NIOEM provides industry-level employments in 277 occupations at the national level. We convert the occupational employment counts into occupational percentages for each industry and map BLS industries to the SIC3 framework.

Even within manufacturing, industries display substantial heterogeneity in their occupational needs. Ellison et al. (2007) calculate a vector correlation of occupational percentages between pairwise industries. Their metric averages 0.47 across the pairs of manufacturing industries, with a range of -0.05 to 1.00. The least correlated industry pair is Logging (241) and Aircrafts and Parts (372) at -0.05. The most correlated industry pair from separate BLS industry groupings is Motor Vehicles and Equipment (371) and Motorcycles, Bicycles, and Parts (375) at 0.98.<sup>11</sup>

This occupational lens allows us to summarize across industries the quality of a city’s labor pool for a new firm in each industry. For a focal industry  $i$  in city  $c$ , we define raw labor suitability as

$$Labor_{ci} = - \sum_{o=1,\dots,O} \left| L_{io} - \left( \sum_{k=1,\dots,I} \frac{E_{kc}}{E_c} L_{ko} \right) \right|, \quad (4)$$

where  $O$  indexes occupations.  $L_{io}$  captures the percentage of industry  $i$ ’s employment in occupation  $o$  taken from the NIOEM. The fraction  $E_{kc}/E_c$  measures the share of city  $c$ ’s incumbent manufacturing employment in industry  $k$ . The internal summation across industries thus interacts the relative composition of SIC3 manufacturing industries in the city with the extent each industry employs the occupation  $o$  in question. This estimated percentage reliance among city  $c$ ’s incumbent manufacturing firms is then differenced from the needs of industry  $i$ . Absolute values of these disparities are summed across 277 occupations to form a metric of the aggregate labor pool suitability. As a final step, this aggregate is multiplied by negative one so that higher values correspond to more suitable labor environments for industry  $i$ ’s startups in city  $c$ .<sup>12</sup>

This metric thus emphasizes the pooled nature of local labor markets. It assumes that it does not matter which manufacturing industries employ local workers, so long as the occupational distribution is suitable for an industry. The metric is by construction orthogonal to city size. The best measured local labor environments are Ship and Boat Building and Repairing’s (373) entry into Bremerton, WA, and Motor Vehicles and Equipment’s (371) entry into Flint, MI. By contrast, the worst labor pool is for Logging’s (241) entry into Flint, MI.

Several limitations of this metric should be noted. First, worker quality is not measured for the local area (nor wage costs). Empirical specifications control for overall differences across

---

<sup>11</sup>Some BLS industries map to two or more SIC3 industries. This multiplicity is not important for the estimations below. Each SIC3 industry is assumed to possess the same occupational composition of employment as that of the BLS industry to which it belongs.

<sup>12</sup>Several examples of (4) can better explain its features. First, if all of city  $c$ ’s manufacturing employment is within industry  $i$  itself, the labor quality will be given a perfect measure of zero for startups in that industry. A perfect market, however, is not exclusive to all firms being in the industry  $i$ . Consider an industry  $i$  that employs equally two occupations  $a$  and  $b$ . A perfect correlation could also descend from city  $c$  having two other incumbent industries that also employ  $a$  and  $b$  in equal proportions. Or, if the incumbent industries are of equal size, one industry may employ  $a$  exclusively and the other  $b$  exclusively. On the other hand, if all of industry  $i$ ’s needs are for a specific occupation that is not employed by any existing industry within city  $c$ , (4) will take its lowest possible value of negative two.

cities in general labor quality but not for unique city-industry quality differences. Second, non-manufacturing industries are not included in the calculation, although some occupations span into other sectors. Endogeneity also remains a clear concern.

### 3.3.3 Technology Spillovers Strength

Startups are also inevitably about new ideas, and the ability of some areas to foster new ideas is one potential reason why they become centers of entrepreneurship. Innovations are rarely created out of wholecloth, but rather come out of intellectual building blocks: new ideas are combinations of old ideas. A third agglomeration economy is the presence of suppliers of ideas, where spatial industrial concentrations fuel entrepreneurship by supporting the transfer of old ideas and the creation of new ones. Marshall (1920) is again the source of this theory, when he emphasized that in industrial clusters "the mysteries of the trade become no mystery, but are, as it were, in the air."

An abundant literature discusses these ideas. Saxenian (1994) illustrates how the flow of ideas helped to create the entrepreneurial cluster of Silicon Valley. Arzaghi and Henderson (2008) document the importance of networking for ad agencies in Manhattan. Glaeser and Kahn (2001) argue that the urbanization of high human-capital industries, like finance, is evidence for the role that density plays in the transfer of ideas, and studies of patent citations highlight the importance of local proximity for scientific exchanges (e.g., Jaffe et al. 1993, Thompson and Fox-Kean 2005, Thompson 2006). Carlino et al. (2007) find that higher urban employment density is correlated with greater patenting per capita within cities, and Rosenthal and Strange (2003) note that intellectual spillovers are strongest at the very local levels of proximity. This concept has further found voice in endogenous growth theory (e.g., Romer 1986, 1990).

To test the importance of local knowledge flows for manufacturing startups, we develop a metric of technology sourcing through patent citations taken from the NBER Patent Database (e.g., Hall et al. 2001). Patent citation patterns can be informative about technology diffusion and knowledge exchanges (e.g., Griliches 1990, Jaffe et al. 2000). We first catalogue the extent to which technologies associated with industry  $i$  cite technologies associated with industry  $k$ , with citation counts being normalized by total citations for each industry.<sup>13</sup> We then calculate  $Tech_{ci}$  for these technology flows that mirrors  $Input_{ci}$  described above. The weakest technology locale is Industrial Inorganic Chemicals (281) in Bremerton, WA. The best measured technology environments are for Ship and Boat Building and Repairing (373) in New London, CT, and Norfolk, VA.

---

<sup>13</sup>We consider over four million citations where both the citing and cited patents are filed within the US after 1975. These citations are then collapsed into a citation matrix using over 400 USPTO technology categories. A probabilistic concordance between the USPTO classification scheme and SIC3 industries is then applied to characterize technology exchanges across industries (e.g., Johnson 1999, Silverman 1999, Kerr 2008a).

Of our three local conditions, intellectual spillovers are the most difficult to quantify and to assess empirically. First, our metric focuses only on technology spillovers. Other intellectual or information spillovers may exist between industries that are not captured by our design, although technology sourcing is a very important form of knowledge sharing in manufacturing. Second, technology flows are not mutually exclusive to the first two Marshallian determinants (e.g., Porter 1990). Technologies embodied in products and machinery can be transferred directly through input-output exchanges. Likewise, industries that share similar labor may also be industries among which there are greater possibilities for intellectual spillovers. Our empirical exercises attempt to isolate technology spillovers by jointly testing  $Tech_{ci}$  with these other two factors, but it is important to note that intellectual spillovers do occur within these channels, too. Endogeneity also remains a concern, as incumbents may locate near startups in hopes of gathering new ideas.<sup>14</sup>

The literature on intellectual spillovers is divided on whether the development of new innovations is most aided by having a large concentration of one’s own industry or by industrial diversity. The view stressing industrial concentration is most often associated with Marshall, Arrow, and Romer (MAR). The MAR model emphasizes the benefits of concentrated industrial centers, particularly citing the gains in increasing returns and learning-by-doing that occur within industries. The second view, often associated with Jacobs (1970), argues that major innovations come when the ideas of one industry are brought into a new industrial sector. This perspective stresses that a wealth of industrial diversity is needed to create the cross-fertilization that leads to new ideas and entrepreneurial success. Duranton and Puga (2001) formalize theoretical foundations for this model.

Recent empirical research has sought to uncover whether industrial specialization or diversity better foster regional development (e.g., Glaeser et al. 1992, Henderson et al. 1995, Feldman and Audretsch 1999). Our accounting for existing city-industry employment mostly captures the MAR model. We also develop a Herfindahl-Hirschman Index (HHI) of employment shares in manufacturing industries for city  $c$  to capture the diversity perspective. The greatest and weakest manufacturing diversities are found in Philadelphia, PA, and Bremerton, WA, respectively. Unlike our other metrics, we do not adjust the HHI metric to have positive coefficients corresponding to support for the Jacobs hypothesis.

In between these two extremes, new empirical research is quantifying the role of related industries in industrial clusters (e.g., Ellison et al. 2007, Kolko 2007, Delgado et al. 2008). This

---

<sup>14</sup>Scherer (1984) develops a technology flow matrix that estimates the extent to which R&D activity in one industry flows out to benefit another industry. This technology transfer occurs either through a supplier-customer relationship between these two industries or through the likelihood that patented inventions obtained in one industry will find applications in the other industry. Patent-based metrics have the advantage of covering the 1975-1999 period, but inventor-to-inventor communication patterns represent a subset of the technology flows encompassed by Scherer (1984). Metrics constructed through the Scherer matrix deliver mostly similar results to those present below.

is precisely what we hope to uncover with our Marshallian and Chinitz connections. We do not designate related industries, however, but instead model how the industrial composition of the city interacts with new startups through the more fundamental channels of goods, people, and ideas.

### 3.4 Entrepreneurial Culture

A final hypothesis is that some areas simply develop a culture of entrepreneurship (e.g., Hofstede 2001). Many observers believe that entrepreneurial hotbeds feed upon themselves, with above-average entry levels encouraging further entry. Saxenian (1994) describes how early entrepreneurs fostered future entrepreneurship in Silicon Valley. Free from the constraints of older industrial structures, these early entrepreneurs developed flat organizational structures, emphasized equity participation by employees, and participated in greater formal and informal cooperation across young startups. These features lead to an industrial structure characterized by vertical disintegration, modular product development, and greater labor mobility across firms (e.g., Fallick et al. 2006). It became the norm, rather than the exception, that employees would strike out on their own, with failure much more respectable than not having tried. Moreover, these attitudes and practices transferred beyond early semiconductor firms to future industries, both related and unrelated.

This sense of entrepreneurial culture is not limited to the exceptional case of Silicon Valley. While outlining his small supplier argument, Chinitz (1961) also noted an "aura of second class citizenship" that surrounds small business owners in cities dominated by big firms. Lamoreaux et al. (2004) describe the strong entrepreneurial culture of Cleveland at the start of the twentieth century, when many startup firms launched the industries of the 2nd Industrial Revolution (e.g., electric light and power, automobiles, chemicals). Florida (2005) emphasizes differences across cities in bohemian cultures and tolerance in the promotion of his creative class. Davidsson (1995) connects differences in entrepreneurial beliefs and new firm formation rates across Swedish regions.

One interpretation of an entrepreneurship culture is that there are agglomeration economies in entrepreneurship. A robust entrepreneurial sector may lead to the development of broader social structures and institutions that support additional entry. For example, angel financiers, specialized educational institutions, and small business lawyers might congregate in areas that start with some initial concentration of entrepreneurship. This clustering may have non-linear externalities similar to the urban capital markets arguments put forward by Helsley and Strange (1991). Further, an agglomeration of entrepreneurs might increase the social returns to taking risks and reduce the stigma associated with entrepreneurial failure. The reduced stigma of failure

is frequently pointed to when discussing differences in entrepreneurial culture across countries (e.g., Landier 2006).

Our approach to assessing this theory is to follow Glaeser (2007) and test whether entrepreneurship in one industry is associated with being located near other industries that are, throughout the US, more entrepreneurial. We examine whether electronics producers are more entrepreneurial when they are located near entrepreneurship-prone industries versus big, vertically-integrated businesses. In other words, is the entrepreneurship of industry  $i$  higher in city  $c$  if other industries in city  $c$  are usually entrepreneurial? The culture metric is

$$Culture_c = \sum_{k=1, \dots, I} \frac{E_{kc} \text{Entry}\%_k^{US}}{E_c \text{Entry}\%^{US}}, \quad (5)$$

where  $\text{Entry}\%_k^{US}$  is entering US establishments divided by existing establishments for industry  $k$ . Higher values of this index indicate that the industry structure surrounding new entrants is characterized by higher expected entry levels based upon US averages. The greatest and weakest cultural measures are for Bremerton, WA, and Charleston, WV, respectively.

There are important limitations with this estimation approach. First, we miss many nuances associated with entrepreneurial culture noted above. We also abstract from ethnic propensities towards starting new enterprises and their uneven spatial distribution (e.g., Fairlie 1999, Kerr 2008b, Wadhwa et al. 2007). Third, the industrial mix of a city may be more exogenous than entrepreneurship levels in those industries in the city, but it is also not entirely exogenous. Places with exogenous characteristics that make entrepreneurship easier should presumably attract more entrepreneurial industries. As such, we interpret these results as, at best, suggestive of the role that entrepreneurial culture might play.

## 4 Empirical Estimations

We now present our empirical estimations of the impact of local conditions on entrepreneurial rates. We first study city-level traits and then turn to conditional estimations that include city and industry fixed effects. We close with estimations that study the entry size distribution and natural cost advantages.

## 4.1 Unconditional City-Industry Estimations

We begin by characterizing city-level traits and entrepreneurship for city  $c$  and industry  $i$ ,

$$\begin{aligned} \ln(Entry_{ci}) = & \eta_i + \beta_P \ln(Pop_c) + \beta_E \ln(Emp_{ci}) + \bar{\gamma}_c \bar{X}_c \\ & + \beta_{Cul} Culture_c + \beta_J Jacobs_c + \beta_C Chinitz_{ci} \\ & + \beta_L Labor_{ci} + \beta_I Input_{ci} + \beta_O Output_{ci} + \beta_T Tech_{ci} + \varepsilon_{ci}, \end{aligned} \tag{6}$$

where  $\eta_i$  is a vector of industry fixed effects that control for fixed differences in industry sizes, entrepreneurship rates, competition, and so on. We further control for city populations and the employment in the city-industry.  $\bar{X}_c$  is a vector of city-level demographics. Variables are transformed to have unit standard deviation to aid interpretation.

The dependent variable  $Entry_{ci}$  is the log measure of mean entry employment by city-industry over 1977 to 1999. We recode less than one entering employee on average as one entering employee for these estimations. This maintains a consistent sample size, and we do not believe that the distinction between zero and one employee at the city-industry level is economically meaningful. Regardless, these cells can be excluded without impacting our results.

We weight estimations by an interaction of the mean industry size within city  $c$  with the mean size of industry  $i$  across cities. This is effectively an interaction of city and industry sizes for urban manufacturing employments. We place more faith in weighted estimations than unweighted estimations since many city-industry observations experience very limited entry. We recognize, however, that weighted estimations may accentuate endogeneity concerns. We thus employ our interaction rather than observed city-industry size. The interaction minimizes the endogeneity spillover for very agglomerated industries, especially in conditional estimations with city and industry fixed effects.

The appendix reports alternative estimations that drop the weights or substitute log entry counts as dependent variables. The emphasized results are mostly robust to these variants, and we indicate below where noticeable differences exist. These differences are typically about the overall elasticities evident in the data, rather than the ordering of explanatory factors.

Table 3 presents our basic results. The first column includes just city populations, city-industry employments, and industry fixed effects. Not surprisingly, both measures of existing agglomeration have strong explanatory power for entrepreneurship rates. Coefficients are interpreted in standard deviations, suggesting for example that a one standard-deviation increase in city size is associated with a half standard-deviation increase in entrepreneurial employment.

The adjusted R-Squared value for this estimation is quite high at 0.80. This explanatory power is primarily due to the existing agglomeration of city-industry employment, which by

itself yields an R-Squared value of 0.66. By contrast, including just industry or city fixed effects yields R-Squared values of 0.29 and 0.47, respectively. This outcome stresses the importance of existing patterns of industrial activity for explaining the spatial distribution of entrepreneurship. As our interests focus on explaining entrepreneurship, versus overall agglomeration, we continue to control for this level. Natural advantages estimations below further relate this predictive power of existing agglomerations to local cost advantages.

The second column incorporates city characteristics: demographics, cultural and diversity measures, and the Chinitz small supplier metric. Somewhat surprisingly, demographics play a very limited role in explaining manufacturing entry patterns. Cities that have a higher population share of young workers, aged 20 to 40 years, tend to have greater manufacturing entry rates (the omitted category is 40 to 60 years in age). We are not able to distinguish whether this gain is due to a greater supply of founders or more suitable workers. Whereas older people are strongly associated with self-employment rates (e.g., Glaeser 2007), they are not associated with new manufacturing startups. This can be partially reconciled through the reduced emphasis of consumption entrepreneurship in manufacturing (e.g., Hurst and Lusardi 2004, Nanda 2008).<sup>15</sup>

A more educated workforce is found to have a negative, but not particularly robust, partial correlation with entry. Unreported estimations that include just the covariates in Column 1 and the bachelors education share find a positive and statistically significant partial correlation for education. Moreover, unweighted regressions do not yield a clear relationship between education levels and entry. We conclude that a robust relationship for education does not exist with our manufacturing entry measure. Future work will hopefully clarify whether this is a particular feature of the manufacturing sector.

In contrast to demographics, the Chinitz measure finds very strong support in these estimations. The coefficient is 0.4 in weighted regressions and 0.1 in unweighted specifications, both being statistically significant and economically important. The Chinitz explanatory power is substantially greater than the culture or industrial diversity metrics. Similar to education, the culture measure is positive and statistically significant when only conditioning on city population and city-industry employment. To some degree, the Chinitz measure may also embody what observers refer to as entrepreneurial culture. The local Herfindahl index metric finds greater specialization supports higher levels of manufacturing entry, even conditional on city-industry employment controls.

Column 3 presents our first evidence on Marshallian agglomeration forces. Each of the factors finds some support in the estimations, but the local labor market mix is found to be

---

<sup>15</sup>Doms et al. (2008) find local skill levels correlate with higher rates of self employment and better startup performance in the US. Bonte et al. (2008) document an inverted-U shape between regional age structures and entrepreneurship rates in Germany.

the most important factor. The full specification (6) is reported in Column 4. Most city-level characteristics remain similar to their individual estimations, and the technology spillover metric remains strong. On the other hand, labor and input suitability measures diminish. In fact, the explanatory power of the Chinitz measure is strong enough to crowd out the general inputs variable of the Marshallian factors completely. We further discuss the Chinitz result below after viewing conditional estimations.

One surprising result from Table 3 was the limited explanatory power of city demographics for manufacturing entrepreneurship. A second surprising feature is the limited explanatory power that additional city traits and Marshallian factors bring. The adjusted R-squared barely increases across the columns to 0.81. In unweighted estimations, the adjusted R-squared hovers around 0.62. This explanatory level also holds in upcoming estimations that control for city fixed effects.

## 4.2 Conditional City-Industry Estimations

Our primary empirical approach looks at variation within cities rather than across them. We test which local industrial attributes are attractive for entrepreneurship. We now turn to estimations that replace the vector of city-level covariates in (6) with a vector of city fixed effects  $\phi_c$ ,

$$\begin{aligned} \ln(\text{Entry}_{ci}) = & \eta_i + \phi_c + \beta_E \ln(\text{Emp}_{ci}) + \beta_C \text{Chinitz}_{ci} \\ & + \beta_L \text{Labor}_{ci} + \beta_I \text{Input}_{ci} + \beta_O \text{Output}_{ci} + \beta_T \text{Tech}_{ci} + \varepsilon_{ci} \end{aligned} \quad (7)$$

This fixed effect estimation removes differences across cities that are common for all industries, for example due to New York’s larger city size. Specifications thus employ within variation: how much of the unexplained city-industry variation in entrepreneurship can we explain through local conditions that are especially suitable for particular industries?

Table 4 presents the basic results. The first regression continues to show that new startups are drawn to existing agglomerations. The combination of initial employment, city fixed effects, and industry fixed effects can together explain 82 percent of the variation in manufacturing entry patterns. In the second regression, we include our three basic agglomeration measures from Marshall (1920). The labor metric continues to be quite strong. A one standard-deviation increase in this variable is associated with a quarter standard-deviation increase in the employment found in new firms. The presence of input suppliers is also important, but industrial customers appear less significant. Workers and suppliers seem to drive location decisions of manufacturing startups, not buyers.

The third regression incorporates the Chinitz measure of small suppliers and again finds a very strong impact on entrepreneurship. A one standard-deviation increase in the presence of

small suppliers correlates with a 0.4 standard-deviation increase in entry levels. As in Table 3, the general inputs metric declines substantially after the Chinitz variable is introduced. The Chinitz variable is also quite important when looking at entry counts. The Chinitz measure is statistically significant in unweighted estimations, but its elasticity declines to around 0.1 and is less differentiated from the Marshallian factors.

While noting the explanatory power of the Chinitz hypothesis, we are cautious about the exceptional strength of this outcome. First, we are quite suspicious about the endogeneity of this measure. The average firm size of supplying industries is surely endogenous to the level of entrepreneurship in the industries they supply. Smaller suppliers may well reflect a smaller, more entrepreneurial customer base. A second concern comes from using firm size to construct the Chinitz metric. To the extent that average firm size reflects local entrepreneurship well, the Chinitz metric may be capturing more a correlation within cities in entrepreneurship levels of related industries rather than a specific supplier relationship.

Encouragingly, the strength of our Chinitz metric parallels two recent findings for the US. Using Dun and Bradstreet Marketplace data, Rosenthal and Strange (2007) note a robust correlation of small establishments in a local area predicting greater entry rates. This correlation holds when looking at differences within metropolitan areas and across multiple industrial sectors. Using Census Bureau data, Drucker and Feser (2007) also find that productivity levels of small plants are reduced when regional industrial concentration increases, although the specific agglomeration channels remain unclear. Further investigation of this topic is an important area for future research.

Column 4 incorporates the technology spillovers metric, again showing general support for Marshall's theories. Labor market pools finds the strongest relative support among the Marshallian factors in weighted estimations, but technology spillovers has greater strength in unweighted specifications and when looking at entry counts. Both of these factors are generally found to be more important than local sales conditions.

Column 5's estimation includes the log employment in facility expansions by existing firms to test whether these local conditions are more important for entrepreneurship specifically versus economic growth more generally. We focus on new facility expansions, versus within-plant employment adjustments, to model discontinuous events that resemble new firm entry (e.g., Kerr and Nanda 2007). Facility expansions are a particularly effective control against unmodeled city policies to promote an industry's growth when such policies are neutral towards startups versus existing firms.

Not surprisingly, high contemporaneous rates of facility expansions predict greater entry by new firms. Most of the other explanatory factors, however, are unaffected by this additional

control. Unreported specifications further incorporate facility expansions in broader SIC2 industry groups. Results are quite similar to those reported, although elasticities for labor pooling sometimes decline. This impact for labor is not too surprising given the substantial occupational sharing across industries within SIC2 groups.

### 4.3 Entry Size Distribution

Table 5 next disaggregates overall entry measures into establishment sizes in the year of entry. The entry of a two-person establishment is presumably a different phenomena than the entry of a new firm with hundreds of employees. We care more about larger entrants in certain contexts, for example when worrying about the determinants of robust local labor demand. On the other hand, the entry of small establishments may be a purer reflection of entrepreneurship and hence more intrinsically interesting. More generally, empirical evidence exists that small and large establishments agglomerate differently (e.g., Holmes and Stevens 2002, Duranton and Overman 2008), and it is useful to extend this description to entering firms.

A second rationale exists for examining the entry size distribution. Better local conditions may foster a larger entry size for entrepreneurs due to factors like less uncertainty about local growth potential and faster assembly of key resources. As discussed in Kerr and Nanda (2007), however, metrics of average entry size confound this intensive margin adjustment with changes in the extensive margin of greater entry rates. Better local conditions may simultaneously foster greater entry by many small firms that leads to an overall decline in average entry size. We feel it is more prudent to look at the distribution measure.

Table 5 finds interesting effects. Existing agglomerations are very important for large entrants, but much less so for small entrants. Weighted estimations tend to find a negative elasticity, while unweighted estimations find a small, positive elasticity. Regardless, the difference in the first row between the smallest-size category and larger-size categories is economically and statistically important. This parallels earlier work on the spatial concentration of the establishment size distribution.

Turning to industrial variables, labor and output variables typically find consistent support throughout the distribution. On the other hand, interesting differences again emerge between the Chinitz and Marshallian inputs measure. The importance of the Chinitz measure declines with entry size, while the importance of general Marshallian conditions strengthens with larger entrants. The Chinitz measure retains an overall higher elasticity throughout Table 5, but in unweighted estimations Marshallian input factors become relatively stronger. Finally, the strength of local technology environments is robustly found to be more important for smaller entrants.

## 4.4 Natural Advantages Estimations

The previous regressions treat city-industry characteristics as exogenous regressors, but surely industrial distributions are themselves endogenous to local entrepreneurship. Reverse causality may play a role if existing firms choose to locate in entrepreneurial hotbeds to interact with startups. It is likewise possible that an unmodeled factor — for example, special industrial policies to encourage the local formation of an industry — are responsible for both startup success and industrial structures present. Both scenarios would bias, usually upward, the parameter estimates and explanatory power. Tables 3 through 5 are best interpreted as partial correlations, showing the connection between new entrants and existing local characteristics.

These endogeneity concerns are perhaps amplified by our calculation of entry rates and industrial structures over the same time horizon in Sections 2 and 3. In these situations, it is tempting to rely on time lags to make causal claims — for example, taking incumbent firms in year  $t - 1$  to predict new entry in year  $t$ . The challenge to this approach, however, is the high persistence in city-industry structures. For example, we find similar outcomes when using pre-1987 industrial structures and entry rates from 1988 onwards. But, we also note that the correlation of city-industry employments from 1976 to 1999 is over 0.8. This persistence limits the potential of time-series approaches for assigning causal directions unless special cases are identified. Even when such "natural experiments" exist, which we discuss further below, they will certainly not be applicable across the whole manufacturing sector.

Instead, we now employ spatial variations in natural cost advantages in an attempt to bound endogeneity issues through reduced-form estimations. Natural spatial distributions of industries, arising from exogenous cost factors outlined in Section 3, provide more robust identification than using actual manufacturing spatial shares. This applies to both own-industry employments and Marshallian linkages across industries. Moreover, grounding our reduced-form estimates in costs differences is very intuitive for manufacturing.

The validity of this approach depends upon the estimation technique. The natural advantages that we model are not strictly exogenous — even coastal access is endogenous over long horizons as waterways shaped city locations for many centuries. While lacking random assignment of city locations, residual variations from conditional estimations are more promising. City fixed effects control for main effects of local natural advantages common across industries, while industry fixed effects control for main effects of industrial technologies and factor uses. Our estimations thus only exploit the residual variation for predicting industrial composition within manufacturing.

We use these residual variations for reduced-form estimators rather than as instruments due to two limitations. First, our cost factors are both mostly pre-determined to the 1976-1999 sam-

ple frame and (at least partially) determined by forces outside of manufacturing. Nevertheless, very persistent agglomerations, like the long history of automobile manufacturing in Detroit, may still shape costs factors and industry intensities that we observe. Second, necessary exclusion restrictions could be violated even if cost factors are strictly exogenous. For example, local industrial policies may build upon local natural advantages.

While not perfect, reduced-form exercises nonetheless build confidence in results found earlier by removing the most worrisome endogeneity. Rather than directly modelling Detroit’s automobile manufacturing employment in 1980 to predict subsequent entry, we are instead using Detroit’s long-run cost factors and how well or poorly they align with the automobile industry’s needs nationally. The reduced-form estimations mirror (7),

$$\begin{aligned} \ln(Entry_{ci}) = & \eta_i + \phi_c + \beta_{NA} \ln(Nat\%_{ci} \cdot Emp_i) \\ & + \beta_L Labor_{ci}^{Nat} + \beta_I Input_{ci}^{Nat} + \beta_O Output_{ci}^{Nat} + \beta_T Tech_{ci}^{Nat} + \varepsilon_{ci}. \end{aligned} \tag{8}$$

We first replace actual city-industry employments with estimated employments from natural spatial distributions and industry sizes. In the case of the existing industry structure, this regression can be interpreted as measuring whether innate cost factors drive entrepreneurship. The predicted log of employment in the city-industry is itself an index of how much local factors favor this particular industry.

In the first column of Table 6, we find that a one standard-deviation increase in predicted city-industry employment increases entry by a full standard deviation. This large coefficient implies that the same cost factors that drive employment levels are also strongly correlated with actual entrepreneurship. Of course, we cannot separate the extent to which this effect reflects the direct impact of lower costs for encouraging entry versus the clustering of startups around initial employments. The adjusted R-Squared value is again quite substantial at 0.77. Regressions that include just predicted employments find an R-Squared of 0.57, compared to 0.66 for actual employments.

In the second column, we incorporate our three measures of Marshallian agglomeration economies. We continue to use national measures of industry interdependencies, but we substitute predicted city-industry distributions for actual city-industry distributions when calculating Marshallian advantages of a local area. As an example, ship building clearly has a natural advantage to be near the coast. But, this attraction further impacts other industries that employ similar workers to ship building, even if these industries do not derive direct benefit for being near the ocean. The coast may also attract suppliers and customers of this industry. Our constructed metrics connect local cost conditions to entrepreneurship through both direct industry advantages and indirect Marshallian interactions.<sup>16</sup>

---

<sup>16</sup> Actual and predicted Marshallian factors have a 0.86 correlation for labor, 0.78 for inputs, 0.56 for outputs, and 0.47 for technology.

Column 2 finds general support for Marshallian factors relating to labor pooling and input-output exchanges. One standard-deviation increases in these indirect linkages descending from natural advantages associate with 0.3 to 0.4 standard-deviation increases in local entry. The output and labor interactions are also robust to many specification variants. Measured elasticities for these two forces decline in unweighted estimations or entry count specifications, but they remain economically and statistically important. The inputs measure, however, is not very robust. Specification variations often yield a negative coefficient that can be statistically significant. Our view is therefore that the natural advantages approach supports the labor pooling and output hypotheses, but it does not support the inputs hypothesis.

Unfortunately, we cannot use the natural advantages approach to confirm the importance of the Chinitz hypothesis. We are unable to predict which suppliers will have big or small firms in particular cities through our current techniques. This is disappointing given the high potential for endogeneity and omitted variables problems with the Chinitz measure. It is also disappointing given the empirical weakness of the general inputs measure. We hope future research will overcome this limitation given the exceptionally strong predictive power of the Chinitz metric in least squares regressions.

The third regression further adds the technology flows metric. A strong negative association is found, but like inputs, this elasticity is not robust to specification variants. Unweighted specifications, for example, find positive associations that Marshallian theory anticipates. Fortunately, coefficients for the other Marshallian factors are robust to including the technology measure. In general, it is very challenging to separate technology and knowledge spillovers from other Marshallian forces. Many technology exchanges within manufacturing closely mirror input-output exchanges or labor similarity. Future research may be able to identify sectors where this overlap is less severe.

Finally, it is noteworthy that our cost advantages approach is stronger for Marshallian factors that operate by pooling across local industries (e.g., labor sharing, output markets) rather than through satisfying a vector of specific needs (e.g., inputs, technology). Exogenous spatial distributions descending from cost advantages will better capture aggregate effects, allowing for above and below expected placements by individual industries in a city, than exact local industrial compositions. Our relative weakness for inputs and technology rationales is likely more due to our blunt metric for vector concepts rather than due to true degrees of economic importance. We hope that future agglomeration theory will outline better pooling versus vector concepts and how best to model them empirically with local cost advantages.

## 4.5 Thoughts for Future Work on Causality

Where do we go from here? Establishing causal relationships between local industrial conditions and entrepreneurship is difficult, but we see several promising paths for future inquiry. First, powerful datasets are emerging with exceptionally detailed firm and employee information. This richness promises more precise assessments than our city-industry approach affords. For example, very detailed breakdowns of plants' material inputs and their sources can better inform the Chinitz hypothesis. Bernard et al. (2008) document extensive product switching by US manufacturing plants, which is a start for tests regarding more general Marshallian input-output relationships. Likewise, linked employer-employee data can characterize the employee-by-employee growth of startups. This mobility data will better assess the overall importance of labor pooling arguments and separate among the various mechanisms noted in Section 3.

The advantages of additional data richness are not confined to more accurate least squares assessments. In the context of specific cities and industries, exogenous shocks to local industrial conditions can be identified. These shocks may descend from corporate restructurings, shifts in local competition, military base closings, and so on. The very rich data becoming available will allow researchers to trace the propagation of specific shocks from directly-affected firms to others through Marshallian channels. For example, import penetration in an industry following trade reforms can have important Marshallian linkages to related firms and industries, even if the imports do not have a direct competitive effect. Greenstone et al. (2007) provide an attractive example through the entry of "million dollar" plants.

A second promising path is to exploit more spatial and temporal variation. We have taken the city to be the relevant unit for economic interactions, but that is clearly a crude approach. Different factors of production have different spatial horizons and markets. Labor pooling is perhaps best modelled through commuting regions given the costs of worker relocation. Input-output linkages extend over longer horizons, with some concave cost relationship likely, except perhaps for the Chinitz arguments. Knowledge transfers may be very localized (e.g., Wall Street) or invariant to distance (e.g., global, high-tech entrepreneurial ties from the US to Israel and Asia). Recent research seeks to more precisely characterize these distances (e.g., Rosenthal and Strange 2003, Ellison et al. 2007).

The agglomeration literature has benefitted from the recent development of continuous geographic measures like Duranton and Overman (2005). We hope that future work will consider local industrial conditions and entry over continuous spatial horizons. The benefits of such an analysis will again extend beyond least squares estimates. The relevant boundaries for Marshallian forces have likely shifted with declines in communications and transportation costs, immigration and mobility changes among the population, and so on. These changes may pro-

vide more causal assessments of how entry follows from Marshallian forces, with the spatial econometrics extending well beyond using lagged local conditions to predict new entry.

## 5 Conclusions

This paper tested a number of hypotheses about determinants of local entrepreneurship rates in manufacturing. We found a very limited role for city demographics or a culture of entrepreneurship in explaining patterns of entry across cities. Likewise, industrial diversity did not promote entry except through the Marshallian channels modelled. The clustering of related industries matter, but not diversity for its own sake. These weak results are not in any sense definitive, as both our metrics are imperfect and our sample is limited to manufacturing. Yet, the results should caution against excessive enthusiasm about these particular explanations for explaining variations in local entrepreneurship levels.

On the other hand, we do find that local costs and other natural advantages variables are very important for new startups. The same natural advantages that predict employment distributions across cities and industries also predict entrepreneurship distributions well. We also used these natural advantage variables to provide us with more exogenous variation in the industrial mix of the city as we tested three main agglomeration theories.

New startups are particularly related to the presence of other industries that hire the same sort of workers. This result holds across all sizes of startups and is independent of weighting strategies. This result also holds in estimations that employ predicted spatial distributions resulting from natural advantages. The broad stability of this finding suggests that people and their human capital are probably the crucial ingredient for most new entrepreneurs.

The evidence on input supplier and customer linkages is somewhat weaker. In general, inputs appear to matter more than customers, and both matter less than the composition of the local labor force. One exception to this finding is that our results strongly support the Chinitz view that small suppliers are critical to entrepreneurship. While this outcome is provocative, we remain cautious about the exceptional strength of this finding. The endogeneity of the Chinitz measure is particularly troubling, since small suppliers are themselves likely to reflect a lot of local entrepreneurship. Our natural advantages approach cannot help with this particular issue, but we hope that future research will uncover other identification strategies for this outcome. There is also a clear need to assess the role of supplier heterogeneity for explaining entry in other sectors.

Overall, these results suggest that local variables do help us to understand the heterogeneity that exists in rates of manufacturing entrepreneurship. Entrepreneurship is itself an important

topic, and we believe that further work in this area can yield high returns. This research should balance comparative analytics, as in this study, with empirical evaluations of specific policy reforms or natural shocks that impact specific local conditions. We hope that our metrics and techniques will aid other researchers in investigations of the manufacturing sector and beyond.

## References

- [1] Acs, Zoltan, and Catherine Armington, *Entrepreneurship, Geography and American Economic Growth* (New York, NY: Cambridge University Press, 2006).
- [2] Acs, Zoltan, Edward Glaeser, Robert Litan, Lee Fleming, Stephan Goetz, William Kerr, Steven Klepper, Stuart Rosenthal, Olav Sorenson, and William Strange, "Entrepreneurship and Urban Success: Toward a Policy Consensus", Kauffman Foundation Report (2008).
- [3] Arzaghi, Mohammad, and J. Vernon Henderson, "Networking off Madison Avenue", forthcoming in *Review of Economic Studies* (2008).
- [4] Audretsch, David, and Maryann Feldman, "R&D Spillovers and the Geography of Innovation and Production", *American Economic Review* 86 (1996), 630-640.
- [5] Autor, David, William Kerr, and Adriana Kugler, "Does Employment Protection Reduce Productivity? Evidence from U.S. States", *Economic Journal* 117:521 (2007), 189-217.
- [6] Bernard, Andrew, Stephen Redding, and Peter Schott, "Multi-Product Firms and Product Switching", Working Paper (2008).
- [7] Blanchflower, David, and Andrew Oswald, "What Makes an Entrepreneur?", *Journal of Labor Economics* 16:1 (1998), 26-60.
- [8] Bönnte, Werner, Oliver Falck, and Stephan Heblich, "The Impact of Regional Age Structure on Entrepreneurship", Working Paper (2008).
- [9] Bozkaya, Ant, and William Kerr, "Labor Regulations and European Industrial Specialization: Evidence from Private Equity Investments", Harvard Business School Working Paper (2007).
- [10] Buenstorf, Guido, and Steven Klepper, "Why Does Entry Cluster Geographically? Evidence from the US Tire Industry", Working Paper (2007).
- [11] Carlino, Gerald, Satyajit Chatterjee, and Robert Hunt, "Urban Density and the Rate of Invention", *Journal of Urban Economics* 61 (2007), 389-419.
- [12] Chinitz, Benjamin, "Contrasts in Agglomeration: New York and Pittsburgh", *American Economic Review* 51:2 (1961), 279-289.
- [13] Combes, Pierre-Philippe, and Gilles Duranton, "Labour Pooling, Labour Poaching, and Spatial Clustering", *Regional Science and Urban Economics* 36 (2006), 1-28.
- [14] Dahl, Michael, and Steven Klepper, "Who Do New Firms Hire?", Working Paper (2007).
- [15] Dahl, Michael, and Olav Sorenson, "Home Sweet Home: Social Capital and Location Choice", Working Paper (2007).
- [16] Davidsson, Per, "Culture, Structure and Regional Levels of Entrepreneurship", *Entrepreneurship & Regional Development* 7:1 (1995), 41-62.
- [17] Davis, Steven, John Haltiwanger, Ron Jarmin, and Javier Miranda, "Volatility and Dispersion in Business Growth Rates: Publicly Traded Versus Privately Held Firms", NBER Working Paper 12354 (2006).

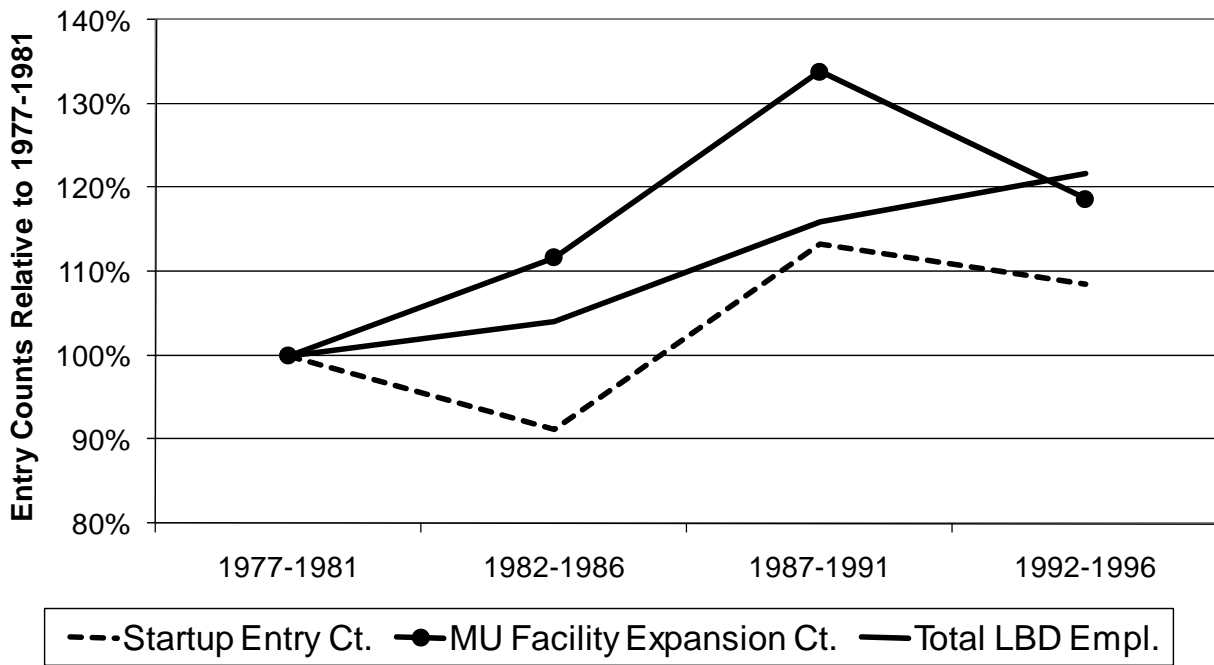
- [18] Davis, Steven, John Haltiwanger, and Scott Schuh, *Job Creation and Destruction* (Cambridge, MA: MIT Press, 1996).
- [19] Delgado, Mercedes, Michael Porter, and Scott Stern, "Convergence, Clusters and Economic Performance", Working Paper (2008).
- [20] Diamond, Charles, and Chris Simon, "Industrial Specialization and the Returns to Labor", *Journal of Labor Economics* 8 (1990), 175-201.
- [21] Doms, Mark, Ethan Lewis, and Alicia Robb, "Local Labor Market Endowments, New Business Characteristics, and Performance", Working Paper (2008).
- [22] Drucker, Joshua, and Edward Feser, "Regional Industrial Dominance, Agglomeration Economies, and Manufacturing Plant Productivity", CES Working Paper 07-31 (2007).
- [23] Dumais, Guy, Glenn Ellison, and Edward Glaeser, "Geographic Concentration as a Dynamic Process", *Review of Economics and Statistics* 84 (2002), 193-204.
- [24] 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.
- [25] 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.
- [26] Duranton, Gilles, and Henry Overman, "Exploring the Detailed Location Patterns of UK Manufacturing Industries Using Microgeographic Data", *Journal of Regional Science* 48:1 (2008), 313-343.
- [27] Duranton, Gilles, and Henry Overman, "Testing for Localization Using Micro-Geographic Data", *Review of Economic Studies* 72 (2005), 1077-1106.
- [28] 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.
- [29] Ellison, Glenn, and Edward Glaeser, "Geographic Concentration in U.S. Manufacturing Industries: A Dartboard Approach", *Journal of Political Economy* 105 (1997), 889-927.
- [30] Ellison, Glenn, and Edward Glaeser, "The Geographic Concentration of Industry: Does Natural Advantage Explain Agglomeration?", *American Economic Review Papers and Proceedings* 89 (1999), 311-316.
- [31] Ellison, Glenn, Edward Glaeser, and William Kerr, "What Causes Industry Agglomeration? Evidence from Coagglomeration Patterns", NBER Working Paper 13068 (2007).
- [32] Evans, David, and Boyan Jovanovic, "An Estimated Model of Entrepreneurial Choice under Liquidity Constraints", *Journal of Political Economy* 97:4 (1989), 808-827.
- [33] Fallick, Bruce, Charles Fleischman, and James Rebitzer, "Job-Hopping in Silicon Valley: Some Evidence Concerning the Microfoundations of a High-Technology Cluster", *Review of Economics and Statistics* 88:3 (2006), 472-481.
- [34] Fairlie, Robert, "The Absence of the African-American Owned Businesses: An Analysis of the Dynamics of Self Employment", *Journal of Labor Economics* 17 (1999), 80-108.

- [35] Feldman, Maryann, "The Locational Dynamics of the US Biotech Industry: Knowledge Externalities and the Anchor Hypothesis", *Industry and Innovation* 10:3 (2003), 311-328.
- [36] Feldman, Maryann, and David Audretsch, "Innovation in Cities: Science-based Diversity, Specialization and Localized Competition", *European Economic Review* 43 (1999), 409-429.
- [37] Figueiredo, Octávio, Paulo Guimaraes, and Douglas Woodward, "Home-Field Advantage: Location Decisions of Portuguese Entrepreneurs", *Journal of Urban Economics* 52:2 (2002), 341-361.
- [38] Florida, Richard, *Cities and the Creative Class* (New York, NY: Routledge, 2005).
- [39] Fujita, Masahisa, Paul Krugman, and Anthony Venables, *The Spatial Economy: Cities, Regions and International Trade* (Cambridge, MA: MIT Press, 1999).
- [40] Glaeser, Edward, "Entrepreneurship and the City", NBER Working Paper 13551 (2007).
- [41] Glaeser, Edward, Hedi Kallal, Jose Scheinkman, and Andrei Shleifer, "Growth in Cities", *Journal of Political Economy* 100 (1992), 1126-1152.
- [42] Glaeser, Edward, and Matthew Kahn, "Decentralized Employment and the Transformation of the American City", NBER Working Paper 8117 (2001).
- [43] Greenstone, Michael, Richard Hornbeck, and Enrico Moretti, "Identifying Agglomeration Spillovers: Evidence from Million Dollar Plants", NBER Working Paper 13833 (2007).
- [44] Griliches, Zvi, "Patent Statistics as Economic Indicators: A Survey", *Journal of Economic Literature* 28 (1990), 1661-1707.
- [45] Hall, Bronwyn, Adam Jaffe, and Manuel Trajtenberg, "The NBER Patent Citation Data File: Lessons, Insights and Methodological Tools", NBER Working Paper 8498 (2001).
- [46] Helsley, Robert, and William Strange, "Agglomeration Economies and Urban Capital Markets", *Journal of Urban Economics* 29 (1991), 96-112.
- [47] Helsley, Robert, and William Strange, "Matching and Agglomeration Economics in a System of Cities", *Regional Science and Urban Economics* 20 (1990), 189-212.
- [48] Henderson, J. Vernon, "Marshall's Scale Economies", *Journal of Urban Economics* 53 (2003), 1-28.
- [49] Henderson, J. Vernon, Ari Kuncoro, and Matt Turner, "Industrial Development in Cities", *Journal of Political Economy* 103:5 (1995), 1067-1090.
- [50] Hofstede, Geert, *Culture and Organizations* (London, UK: Harper Collins, 2001).
- [51] Holmes, Thomas, and John Stevens, "Geographic Concentration and Establishment Scale", *Review of Economics and Statistics* 84:4 (2002), 682-690.
- [52] Hurst, Erik, and Annamaria Lusardi, "Liquidity Constraints, Household Wealth, and Entrepreneurship", *Journal of Political Economy* 112:2 (2004), 319-347.
- [53] Jacobs, Jane, *The Economy of Cities* (New York, NY: Vintage Books, 1970).

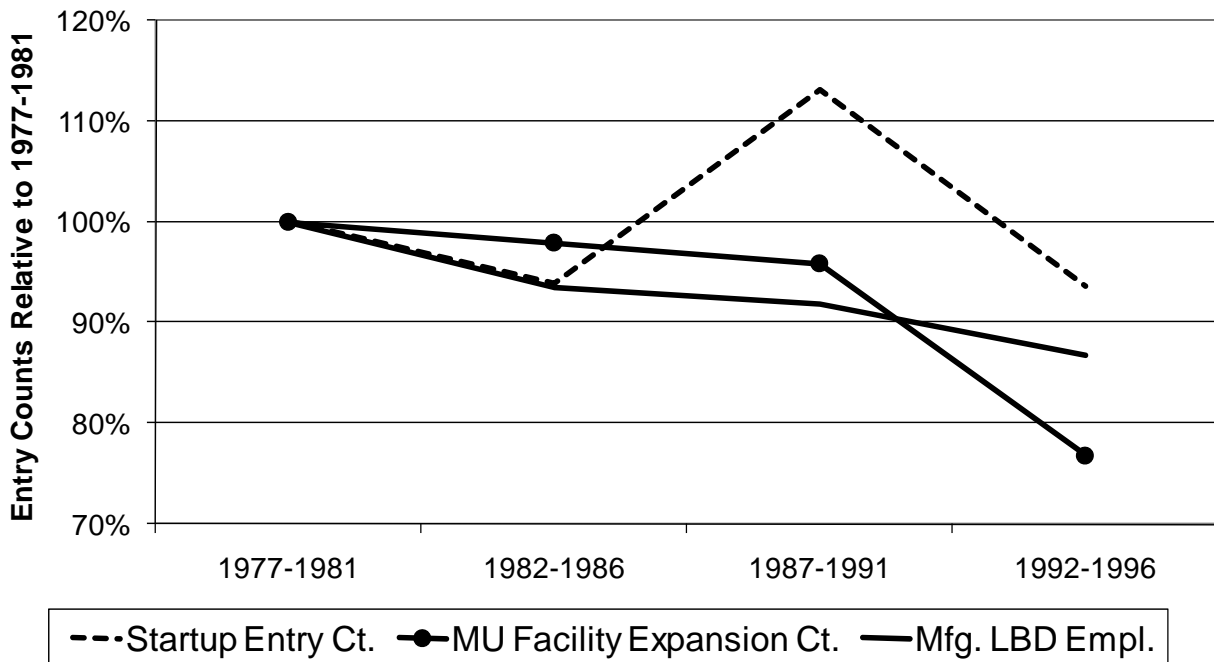
- [54] Jaffe, Adam, Manuel Trajtenberg, and Michael Fogarty, "Knowledge Spillovers and Patent Citations: Evidence from a Survey of Inventors", *American Economic Review Papers and Proceedings* 90 (2000), 215-218.
- [55] Jaffe, Adam, Manuel Trajtenberg, and Rebecca Henderson, "Geographic Localization of Knowledge Spillovers as Evidenced by Patent Citations", *The Quarterly Journal of Economics* 108:3 (1993), 577-598.
- [56] Jarmin, Ron, and Javier Miranda, "The Longitudinal Business Database", CES Working Paper (2002).
- [57] Johnson, Daniel, "150 Years of American Invention: Methodology and a First Geographic Application", Wellesley College Economics Working Paper 99-01 (1999).
- [58] Kerr, William, "Ethnic Scientific Communities and International Technology Diffusion", *Review of Economics and Statistics* 90 (2008a), 518-537.
- [59] Kerr, William, "The Agglomeration of US Ethnic Inventors", Harvard Business School Working Paper 09-003 (2008b).
- [60] Kerr, William, and Ramana Nanda, "Democratizing Entry: Banking Deregulation, Financing Constraints, and Entrepreneurship", Harvard Business School Working Paper 07-033 (2007).
- [61] Kim, Sukkoo, "Regions, Resources, and Economic Geography: Sources of U.S. Regional Comparative Advantage, 1880-1987", *Regional Science and Urban Economics* 29:1 (1999), 123-137.
- [62] Kolko, Jed, "Agglomeration and Co-Agglomeration of Services Industries", Working Paper (2007).
- [63] Krugman, Paul, *Geography and Trade* (Cambridge, MA: MIT Press, 1991).
- [64] Lamoreaux, Naomi, Margaret Levenstein, and Kenneth Sokoloff, "Financing Invention During the Second Industrial Revolution: Cleveland, Ohio, 1870-1920", NBER Working Paper 10923 (2004).
- [65] Landier, Augustin, "Entrepreneurship and the Stigma of Failure", Working Paper (2006).
- [66] Marshall, Alfred, *Principles of Economics* (London, UK: MacMillan and Co., 1920).
- [67] Marx, Matt, Deborah Strumsky, and Lee Fleming, "Noncompetes and Inventor Mobility: Specialists, Stars, and the Michigan Experiment", Harvard Business School Working Paper 07-042 (2007).
- [68] McGuckin, Robert, and Suzanne Peck, "Manufacturing Establishments Reclassified into New Industries: The Effect of Survey Design Rules", CES Working Paper (1992).
- [69] Michelacci, Claudio, and Olmo Silva, "Why So Many Local Entrepreneurs?", *Review of Economics and Statistics* 89:4 (2007), 615-633.
- [70] Nanda, Ramana, "Cost of External Finance and Selection into Entrepreneurship", Harvard Business School Working Paper (2008).
- [71] Overman, Henry, and Diego Puga, "Labor Pooling as a Source of Agglomeration: An Empirical Investigation", Working Paper (2007).

- [72] Porter, Michael, *The Competitive Advantage of Nations* (New York, NY: The Free Press, 1990).
- [73] Romer, Paul, "Increasing Returns and Long-Run Growth", *Journal of Political Economy* 94:5 (1986), 1002-1037.
- [74] Romer, Paul, "Endogenous Technological Change", *Journal of Political Economy* 98:5 (1990), S71-S102.
- [75] Rosenthal, Stuart, and William Strange, "Geography, Industrial Organization, and Agglomeration", *Review of Economics and Statistics* 85:2 (2003), 377-393.
- [76] Rosenthal, Stuart, and William Strange, "Small Establishments/Big Effects: Agglomeration, Industrial Organization and Entrepreneurship", Working Paper (2007).
- [77] Rosenthal, Stuart, and William Strange, "The Determinants of Agglomeration", *Journal of Urban Economics* 50 (2001), 191-229.
- [78] Rotemberg, Julio, and Garth Saloner, "Competition and Human Capital Accumulation: A Theory of Interregional Specialization and Trade", *Regional Science and Urban Economics* 30 (2000), 373-404.
- [79] Saxenian, AnnaLee, *Regional Advantage: Culture and Competition in Silicon Valley and Route 128* (Cambridge, MA: Harvard University Press, 1994).
- [80] Scherer, Frederic M., "Using Linked Patent Data and R&D Data to Measure Technology Flows", in Griliches, Zvi (ed.) *R & D, Patents and Productivity* (Chicago, IL: The University of Chicago Press, 1984).
- [81] Silverman, Brian, "Technological Resources and the Direction of Corporate Diversification: Toward an Integration of the Resource-Based View and Transaction Cost Economics", *Management Science* 45 (1999), 1109-1124.
- [82] Thompson, Peter, "Patent Citations and the Geography of Knowledge Spillovers: Evidence from Inventor- and Examiner-Added Citations", *Review of Economics and Statistics* 88:2 (2006), 383-388.
- [83] Thompson, Peter, and Melanie Fox-Kean, "Patent Citations and the Geography of Knowledge Spillovers: A Reassessment", *American Economic Review* 95:1 (2005), 450-460.
- [84] Wadhwa, Vivek, AnnaLee Saxenian, Ben Rissing, and Gary Gereffi, "America's New Immigrant Entrepreneurs I", Working Paper (2007).
- [85] Zucker, Lynne, Michael Darby, and Marilyn Brewer, "Intellectual Human Capital and the Birth of U.S. Biotechnology Enterprises", *American Economic Review* 88 (1998), 290-306.

**Fig. 1a: Relative Birth Counts**  
All Sectors, By Type of Entrant

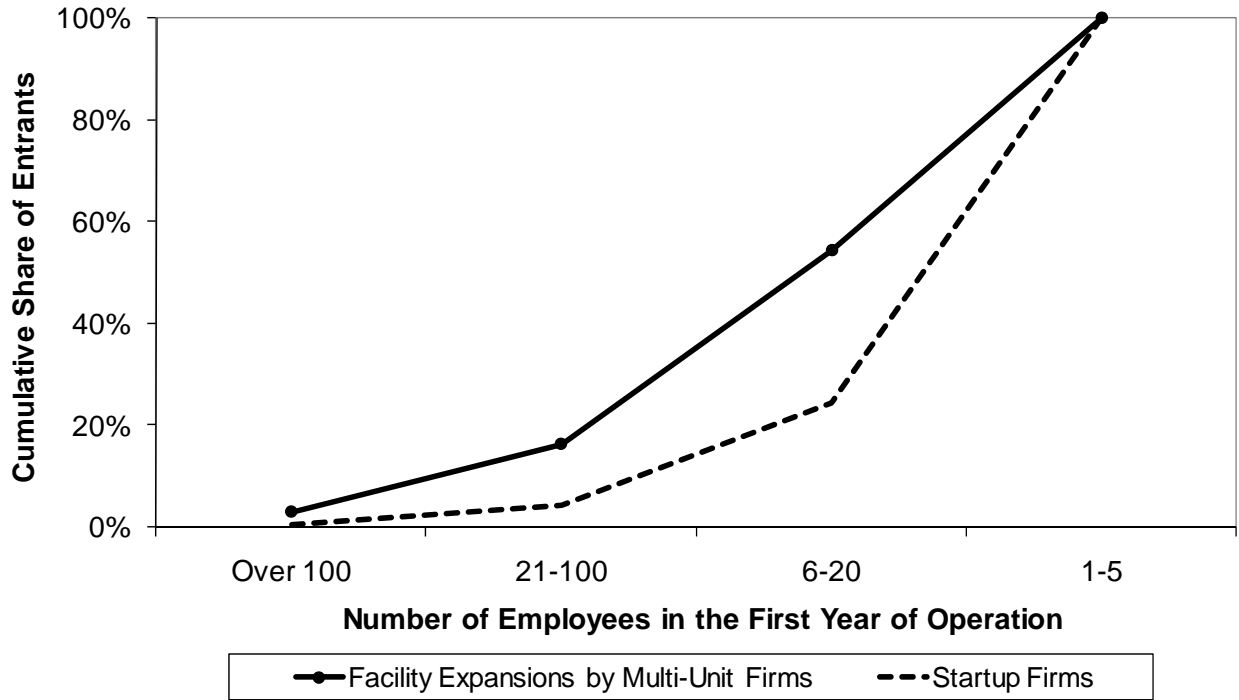


**Fig. 1b: Relative Birth Counts**  
Manufacturing Sector, By Type of Entrant



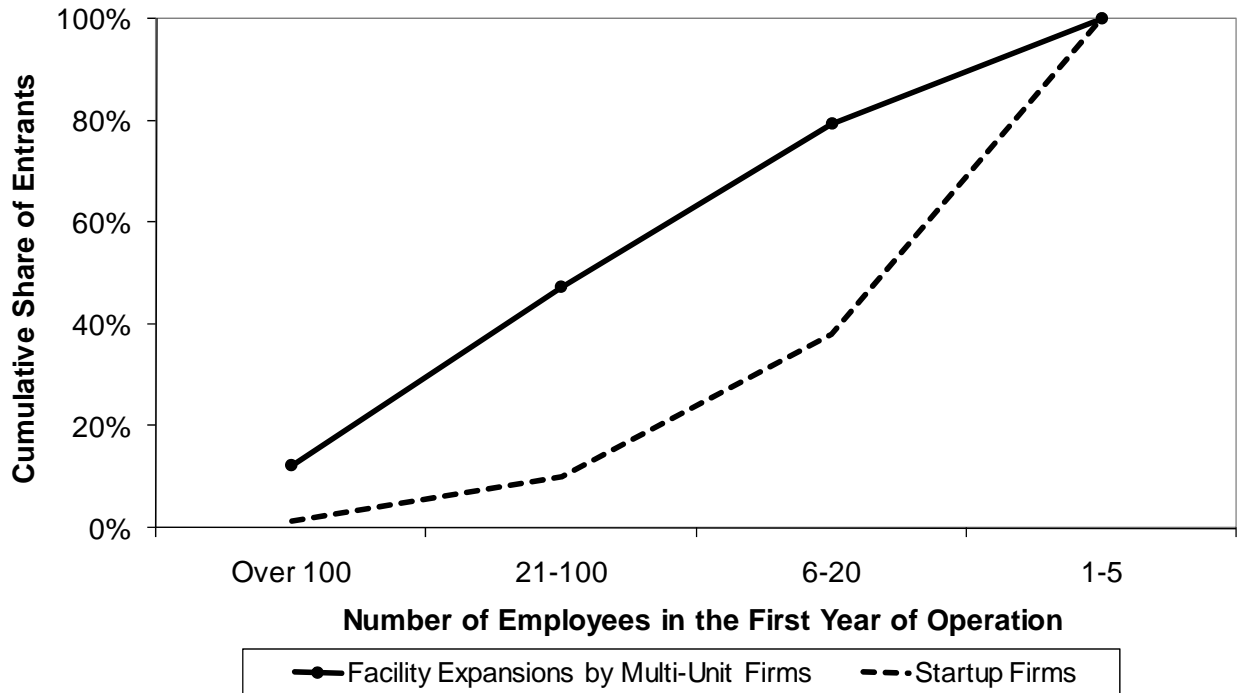
### Fig. 2a: Distribution of Entry Sizes

All Sectors, By Type of Entrant



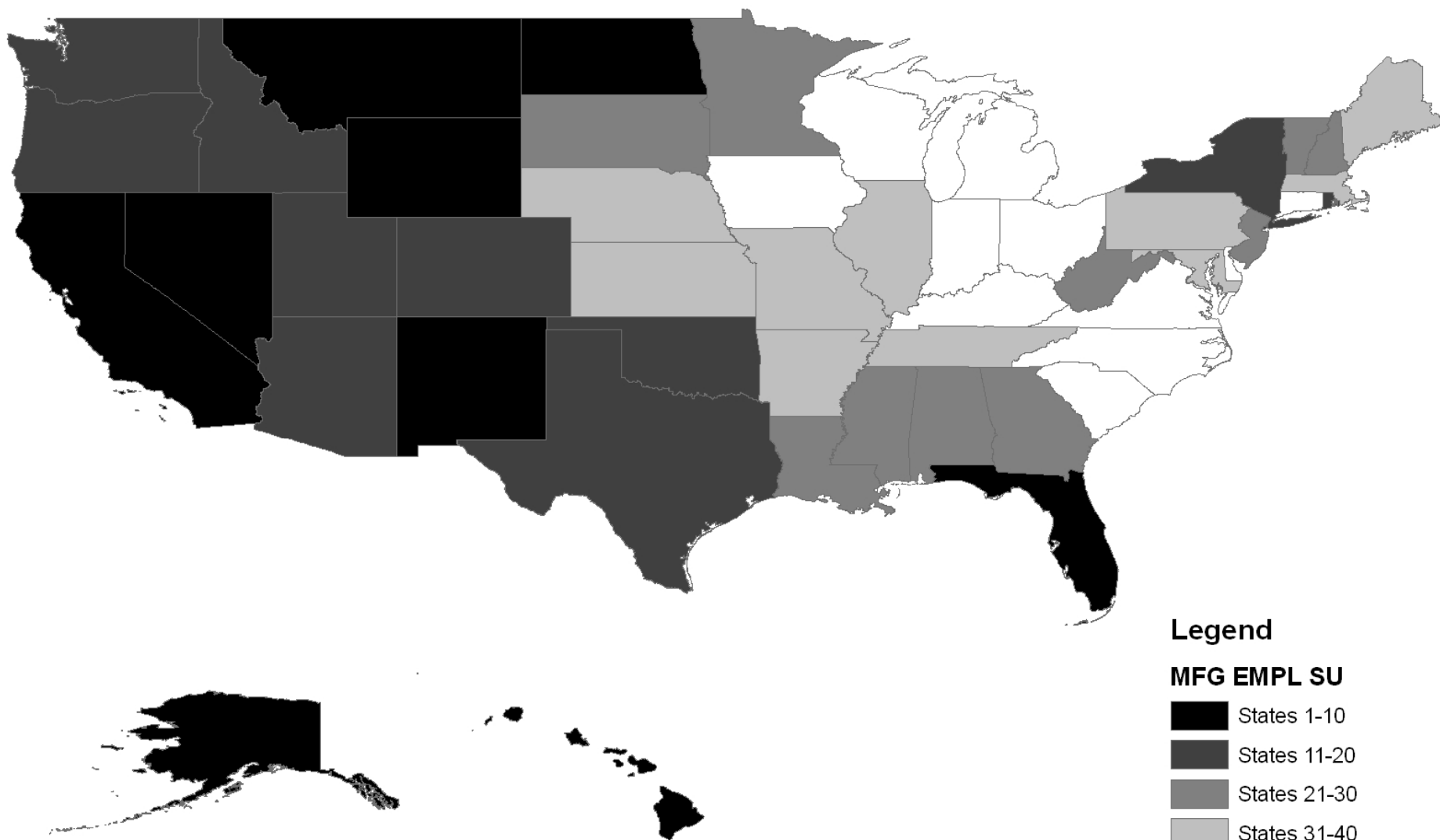
### Fig. 2b: Distribution of Entry Sizes

Manufacturing Sector, By Type of Entrant



# Fig. 3a: Spatial Distribution of Mfg. Entry

States Ranked by Employment Share of Entering Firms



## Legend

### MFG EMPL SU

- States 1-10
- States 11-20
- States 21-30
- States 31-40
- States 41-50

# Fig. 3b: Spatial Distribution of Mfg. Entry

States Ranked by Count Share of Entering Firms

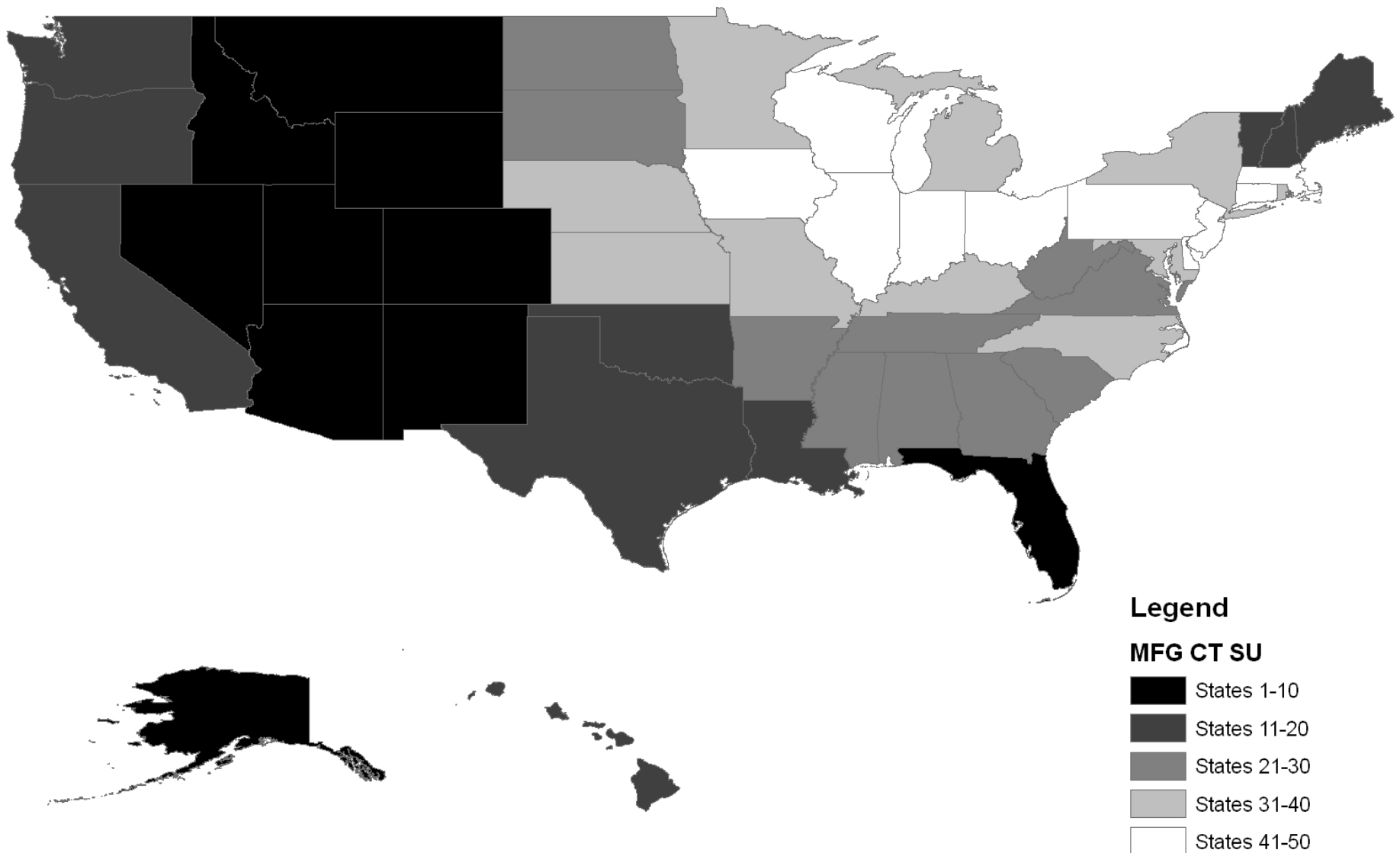


Table 1: LBD Descriptive Statistics on US Entry Rates

	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: Local Industrial Conditions for Mfg. Entry

	Mean	Standard Deviation
<i>A. Actual LBD Industrial Conditions</i>		
Share of Industry in a City	0.004	0.014
Total Mfg. Empl. in City-Industry	353.7	1594.2
Startup Mfg. Empl. in City-Industry	6.0	26.0
Total Mfg. Firms in City-Industry	6.4	28.1
Startup Mfg. Firms in City-Industry	0.5	2.3
Labor Market Strength	-0.825	0.202
Input / Supplier Strength	-1.185	0.136
Chinitz Index of Small Suppliers (x1000)	0.173	0.212
Output / Customer Strength (x1000)	2.735	6.509
Technology Strength - Patents	-1.412	0.209
Cultural Measure of Pred. Entry Rate	0.991	0.099
HHI Index of Mfg. Employment in City	0.088	0.088
Share - Bachelors Education and Above	0.199	0.065
Share - Age 19 and Younger	0.291	0.032
Share - Age 20 to 39	0.331	0.038
Share - Age 60 and Older	0.166	0.042
<i>B. Predicted Industrial Conditions using Natural Advantages</i>		
Share of Industry in a City	0.004	0.006
Total Mfg. Empl. in City-Industry	353.4	956.9
Labor Market Strength	-0.792	0.195
Input / Supplier Strength	-1.102	0.143
Output / Customer Strength (x1000)	2.757	4.181
Technology Strength - Patents	-1.026	0.195

Notes: All pairwise combinations of manufacturing SIC3 industries and cities are included, except those listed in the text, for 33,550 observations. Metrics combine LBD industrial structures within a city with industry traits to describe strength of local industrial conditions. LBD employment and firm counts are annual means for 1977-1999. Labor indices are calculated from the BLS National Industry-Occupation Employment Matrix for 1987. Input-Output relationships are calculated from the BEA Benchmark Input-Output Matrix for 1987. Technology Flows are calculated from the NBER Patent Citation Database for 1975-1997. Demographics are calculated from the 1990 Census of Populations MSA Sample. Variables are transformed from these raw values to have unit standard deviation before estimations.

Table 3: Unconditional Estimations of Mfg. Entrepreneurship Rates

	Base Estimation	City Traits	Marshallian Factors	Full Estimation
	(1)	(2)	(3)	(4)
<i>DV is Log Entry Empl. in New Firms by City-Industry</i>				
Log of City Population	0.499 (0.013)	0.511 (0.013)	0.530 (0.014)	0.516 (0.013)
Log of Employment in City-Industry	0.720 (0.020)	0.610 (0.017)	0.586 (0.018)	0.566 (0.017)
<u>City Characteristics:</u>				
Share of Population with Bachelors Ed.		-0.067 (0.017)		-0.084 (0.017)
Share of Population under 20 in Age		0.021 (0.018)		0.005 (0.018)
Share of Population 20-40 in Age		0.146 (0.028)		0.130 (0.028)
Share of Population over 60 in Age		0.006 (0.027)		-0.014 (0.027)
Cultural Metric of Predicted Entr. Rates		-0.017 (0.012)		-0.010 (0.011)
HHI Index of Mfg. Employment in City		0.025 (0.009)		0.039 (0.010)
Chinitz Measure of Small Suppliers		0.401 (0.020)		0.378 (0.021)
<u>Marshallian Factors:</u>				
Labor Market Strength Metric			0.184 (0.027)	0.047 (0.025)
Inputs / Supplier Strength Metric			0.044 (0.025)	-0.056 (0.025)
Outputs / Customer Strength Metric			0.018 (0.014)	0.022 (0.012)
Technology Strength Metric			0.054 (0.016)	0.079 (0.016)
Adjusted R-Squared	0.80	0.81	0.80	0.81
SIC3 Fixed Effects	X	X	X	X

Notes: Estimations consider log entry employments of new firms for city-industries taken from the LBD. Entry employments are annual averages for city-industries over the 1977-1999 period. Zero employment is recoded as a single employment for these estimations. The construction of the independent regressors is described in the text. Estimations report robust standard errors and have 33,550 observations. Weighted regressions employ an interaction of average industry size across cities with average size of industries within a city. Variables are transformed to have unit standard deviation for interpretation.

Table 4: Conditional Estimations of Mfg. Entrepreneurship Rates

	Existing Industry Structure	Add Labor + I/O Flow	Add Chinitz Metric	Add Technology Flows	Add Facility Expansions
	(1)	(2)	(3)	(4)	(5)
<i>DV is Log Entry Empl. in New Firms by City-Industry</i>					
Log of Employment in City-Industry	0.663 (0.020)	0.514 (0.018)	0.478 (0.018)	0.471 (0.018)	0.349 (0.018)
Labor Market Strength Metric		0.240 (0.028)	0.128 (0.025)	0.097 (0.027)	0.083 (0.027)
Inputs / Supplier Strength Metric		0.119 (0.028)	-0.023 (0.029)	-0.038 (0.028)	-0.076 (0.028)
Chinitz Measure of Small Suppliers			0.386 (0.026)	0.382 (0.026)	0.321 (0.025)
Outputs / Customer Strength Metric		0.017 (0.014)	0.036 (0.012)	0.031 (0.012)	0.020 (0.012)
Technology Strength Metric				0.077 (0.021)	0.052 (0.021)
Log Facility Entry by Multi-Unit Firms					0.182 (0.011)
Adjusted R-Squared	0.82	0.83	0.83	0.83	0.85
City Fixed Effects	X	X	X	X	X
SIC3 Fixed Effects	X	X	X	X	X

Notes: Estimations consider log entry employments in new firms for city-industries taken from the LBD. Entry employments are annual averages for city-industries over the 1977-1999 period. Zero employment is recoded as a single employment for these estimations. The construction of the independent regressors is described in the text. Estimations report robust standard errors and have 33,550 observations. Weighted regressions employ an interaction of average industry size across cities with average size of industries within a city. Variables are transformed to have unit standard deviation for interpretation.

Table 5: Entry Size Distribution Estimations

	Total	Entering Employment Of			
	Entry	1-5	6-20	21-100	101+
	(1)	(2)	(3)	(4)	(5)
<i>DV is Log Entry Empl. in New Firms by City-Industry</i>					
Log of Employment in City-Industry	0.471 (0.018)	-0.052 (0.029)	0.138 (0.022)	0.297 (0.025)	0.420 (0.038)
Labor Market Strength Metric	0.097 (0.027)	0.205 (0.055)	0.165 (0.038)	0.136 (0.041)	0.163 (0.075)
Inputs / Supplier Strength Metric	-0.038 (0.028)	-0.141 (0.058)	-0.089 (0.039)	-0.066 (0.046)	0.143 (0.070)
Chinitz Measure of Small Suppliers	0.382 (0.026)	0.805 (0.059)	0.643 (0.043)	0.583 (0.043)	0.510 (0.056)
Outputs / Customer Strength Metric	0.031 (0.012)	0.039 (0.024)	0.042 (0.020)	0.059 (0.018)	0.061 (0.033)
Technology Strength Metric	0.077 (0.021)	0.314 (0.044)	0.222 (0.032)	0.229 (0.033)	0.067 (0.058)
Adjusted R-Squared	0.83	0.82	0.84	0.77	0.55
City Fixed Effects	X	X	X	X	X
SIC3 Fixed Effects	X	X	X	X	X

Notes: See Table 4. Estimations disaggregate entry into an entry size distribution based upon initial employment in the establishment.

Table 6: Natural Cost Advantages Estimations

	Existing Industry Structure	Add Labor + I/O Flow	Add Technology Flows
	(1)	(2)	(3)
<i>DV is Log Entry Empl. in New Firms by City-Industry</i>			
Pred. Log of Empl. in City-Industry	1.000 (0.130)	0.515 (0.124)	0.667 (0.140)
Pred. Labor Market Strength Metric		0.384 (0.155)	0.495 (0.165)
Pred. Inputs / Supplier Strength Metric		0.273 (0.135)	0.309 (0.135)
Pred. Outputs / Customer Strength		0.345 (0.126)	0.373 (0.126)
Pred. Technology Strength Metric			-0.312 (0.134)
Adjusted R-Squared	0.77	0.77	0.77
City Fixed Effects	X	X	X
SIC3 Fixed Effects	X	X	X

Notes: See Table 4. Estimations replace actual industry distributions for calculating explanatory variables with predicted industry distributions from natural cost advantages.

App. Table 1: Unconditional Estimations of Mfg. Entry Employments

	Unweighted Estimations				Weighted Estimations			
	Base	City	Marsh.	Full	Base	City	Marsh.	Full
	Est.	Traits	Factors	Est.	Est.	Traits	Factors	Est.
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
<i>Dep. Variable is Log Entry Employment in New Firms by City-Industry</i>								
Log of City Population	0.304 (0.005)	0.306 (0.005)	0.306 (0.006)	0.301 (0.006)	0.499 (0.013)	0.511 (0.013)	0.530 (0.014)	0.516 (0.013)
Log of Employment in City-Industry	0.448 (0.006)	0.431 (0.006)	0.404 (0.006)	0.401 (0.006)	0.720 (0.020)	0.610 (0.017)	0.586 (0.018)	0.566 (0.017)
<u>City Characteristics:</u>								
Share of Population with Bachelors Ed.		0.012 (0.005)		0.010 (0.005)		-0.067 (0.017)		-0.084 (0.017)
Share of Population under 20 in Age		-0.002 (0.006)		-0.004 (0.006)		0.021 (0.018)		0.005 (0.018)
Share of Population 20-40 in Age		0.008 (0.007)		0.005 (0.007)		0.146 (0.028)		0.130 (0.028)
Share of Population over 60 in Age		-0.007 (0.008)		-0.012 (0.009)		0.006 (0.027)		-0.014 (0.027)
Cultural Metric of Predicted Entr. Rates		0.003 (0.004)		0.005 (0.004)		-0.017 (0.012)		-0.010 (0.011)
HHI Index of Mfg. Employment in City		0.009 (0.003)		0.021 (0.004)		0.025 (0.009)		0.039 (0.010)
Chinitz Measure of Small Suppliers		0.101 (0.007)		0.092 (0.007)		0.401 (0.020)		0.378 (0.021)
<u>Marshallian Factors:</u>								
Labor Market Strength Metric			0.040 (0.008)	0.018 (0.008)			0.184 (0.027)	0.047 (0.025)
Inputs / Supplier Strength Metric			0.032 (0.008)	0.005 (0.009)			0.044 (0.025)	-0.056 (0.025)
Outputs / Customer Strength Metric			0.045 (0.008)	0.032 (0.007)			0.018 (0.014)	0.022 (0.012)
Technology Strength Metric			0.023 (0.006)	0.042 (0.007)			0.054 (0.016)	0.079 (0.016)
Adjusted R-Squared	0.62	0.63	0.62	0.63	0.80	0.81	0.80	0.81
SIC3 Fixed Effects	X	X	X	X	X	X	X	X

Notes: See Table 3.

App. Table 2: Conditional Estimations of Mfg. Entry Employments

	Unweighted Estimations					Weighted Estimations				
	Existing Industry Structure	Add Labor + I/O Flow	Add Chinitz Metric	Add Tech Flows	Add Facility Expansions	Existing Industry Structure	Add Labor + I/O Flow	Add Chinitz Metric	Add Tech Flows	Add Facility Expansions
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
<i>Dep. Variable is Log Entry Employment in New Firms by City-Industry</i>										
Log of Employment in City-Industry	0.428 (0.006)	0.372 (0.006)	0.366 (0.006)	0.360 (0.006)	0.259 (0.006)	0.663 (0.020)	0.514 (0.018)	0.478 (0.018)	0.471 (0.018)	0.349 (0.018)
Labor Market Strength Metric		0.065 (0.009)	0.054 (0.009)	0.030 (0.009)	0.024 (0.009)		0.240 (0.028)	0.128 (0.025)	0.097 (0.027)	0.083 (0.027)
Inputs / Supplier Strength Metric		0.092 (0.009)	0.052 (0.010)	0.044 (0.010)	0.011 (0.009)		0.119 (0.028)	-0.023 (0.029)	-0.038 (0.028)	-0.076 (0.028)
Chinitz Measure of Small Suppliers			0.080 (0.008)	0.080 (0.008)	0.072 (0.007)			0.386 (0.026)	0.382 (0.026)	0.321 (0.025)
Outputs / Customer Strength Metric		0.042 (0.008)	0.036 (0.007)	0.032 (0.007)	0.018 (0.007)		0.017 (0.014)	0.036 (0.012)	0.031 (0.012)	0.020 (0.012)
Technology Strength Metric				0.061 (0.010)	0.044 (0.009)				0.077 (0.021)	0.052 (0.021)
Log Facility Entry by Multi-Unit Firms					0.239 (0.007)					0.182 (0.011)
Adjusted R-Squared	0.65	0.66	0.66	0.66	0.69	0.82	0.83	0.83	0.83	0.85
City Fixed Effects	X	X	X	X	X	X	X	X	X	X
SIC3 Fixed Effects	X	X	X	X	X	X	X	X	X	X

Notes: See Table 4.

App. Table 3: Conditional Estimations of Mfg. Entry Employments by Entrant Size

	Unweighted Estimations					Weighted Estimations				
	Total	Entering Employment Of				Total	Entering Employment Of			
	Entry	1-5	6-20	21-100	101+	Entry	1-5	6-20	21-100	101+
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
<i>Dep. Variable is Log Entry Employment in New Firms by City-Industry</i>										
Log of Employment in City-Industry	0.360 (0.006)	0.059 (0.005)	0.139 (0.006)	0.239 (0.006)	0.259 (0.009)	0.471 (0.018)	-0.052 (0.029)	0.138 (0.022)	0.297 (0.025)	0.420 (0.038)
Labor Market Strength Metric	0.030 (0.009)	0.021 (0.010)	0.026 (0.009)	0.026 (0.010)	0.043 (0.013)	0.097 (0.027)	0.205 (0.055)	0.165 (0.038)	0.136 (0.041)	0.163 (0.075)
Inputs / Supplier Strength Metric	0.044 (0.010)	0.015 (0.012)	0.028 (0.011)	0.056 (0.011)	0.126 (0.014)	-0.038 (0.028)	-0.141 (0.058)	-0.089 (0.039)	-0.066 (0.046)	0.143 (0.070)
Chinitz Measure of Small Suppliers	0.080 (0.008)	0.114 (0.016)	0.114 (0.012)	0.082 (0.008)	0.021 (0.009)	0.382 (0.026)	0.805 (0.059)	0.643 (0.043)	0.583 (0.043)	0.510 (0.056)
Outputs / Customer Strength Metric	0.032 (0.007)	0.026 (0.009)	0.026 (0.008)	0.027 (0.008)	0.053 (0.012)	0.031 (0.012)	0.039 (0.024)	0.042 (0.020)	0.059 (0.018)	0.061 (0.033)
Technology Strength Metric	0.061 (0.010)	0.114 (0.012)	0.097 (0.011)	0.091 (0.012)	0.037 (0.015)	0.077 (0.021)	0.314 (0.044)	0.222 (0.032)	0.229 (0.033)	0.067 (0.058)
Adjusted R-Squared	0.66	0.53	0.60	0.52	0.29	0.83	0.82	0.84	0.77	0.55
City Fixed Effects	X	X	X	X	X	X	X	X	X	X
SIC3 Fixed Effects	X	X	X	X	X	X	X	X	X	X

Notes: See Table 5.

App. Table 4: Conditional Estimations of Mfg. Entry Employments through Natural Advantages

	Unweighted Estimations			Weighted Estimations		
	Existing Industry Structure	Add Labor + I/O Flow	Add Tech Flows	Existing Industry Structure	Add Labor + I/O Flow	Add Tech Flows
	(1)	(2)	(3)	(4)	(5)	(6)
<i>Dep. Variable is Log Entry Employment in New Firms by City-Industry</i>						
Pred. Log of Empl. in City-Industry	0.424 (0.031)	0.402 (0.032)	0.389 (0.033)	1.000 (0.130)	0.515 (0.124)	0.667 (0.140)
Pred. Labor Market Strength Metric		0.089 (0.049)	0.071 (0.052)		0.384 (0.155)	0.495 (0.165)
Pred. Inputs / Supplier Strength Metric		-0.146 (0.050)	-0.144 (0.050)		0.273 (0.135)	0.309 (0.135)
Pred. Outputs / Customer Strength		0.054 (0.028)	0.054 (0.028)		0.345 (0.126)	0.373 (0.126)
Pred. Technology Strength Metric			0.053 (0.044)			-0.312 (0.134)
Adjusted R-Squared	0.57	0.58	0.58	0.77	0.77	0.77
City Fixed Effects	X	X	X	X	X	X
SIC3 Fixed Effects	X	X	X	X	X	X

Notes: See Table 6.

App. Table 5: Unconditional Estimations of Mfg. Entry Counts

	Unweighted Estimations				Weighted Estimations			
	Base	City	Marsh.	Full	Base	City	Marsh.	Full
	Est.	Traits	Factors	Est.	Est.	Traits	Factors	Est.
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
<i>Dep. Variable is Log Entry Count of New Firms by City-Industry</i>								
Log of City Population	0.318 (0.008)	0.323 (0.008)	0.333 (0.009)	0.326 (0.009)	0.998 (0.039)	1.038 (0.039)	1.100 (0.040)	1.068 (0.040)
Log of Employment in City-Industry	0.130 (0.006)	0.110 (0.005)	0.090 (0.006)	0.085 (0.006)	0.474 (0.047)	0.226 (0.037)	0.180 (0.038)	0.120 (0.035)
<u>City Characteristics:</u>								
Share of Population with Bachelors Ed.		0.005 (0.005)		0.003 (0.005)		-0.311 (0.053)		-0.355 (0.052)
Share of Population under 20 in Age		-0.017 (0.007)		-0.022 (0.007)		0.028 (0.042)		-0.020 (0.043)
Share of Population 20-40 in Age		0.016 (0.008)		0.010 (0.008)		0.503 (0.086)		0.457 (0.088)
Share of Population over 60 in Age		-0.018 (0.010)		-0.026 (0.010)		0.018 (0.064)		-0.040 (0.065)
Cultural Metric of Predicted Entr. Rates		0.001 (0.004)		0.002 (0.004)		0.020 (0.027)		0.030 (0.026)
HHI Index of Mfg. Employment in City		0.027 (0.004)		0.035 (0.004)		0.127 (0.024)		0.140 (0.026)
Chinitz Measure of Small Suppliers		0.125 (0.014)		0.116 (0.014)		0.967 (0.059)		0.905 (0.060)
<u>Marshallian Factors:</u>								
Labor Market Strength Metric			0.069 (0.010)	0.044 (0.010)			0.566 (0.072)	0.224 (0.064)
Inputs / Supplier Strength Metric			0.014 (0.010)	-0.022 (0.011)			0.025 (0.074)	-0.197 (0.073)
Outputs / Customer Strength Metric			0.057 (0.012)	0.041 (0.010)			0.080 (0.044)	0.077 (0.038)
Technology Strength Metric			-0.004 (0.008)	0.025 (0.009)			0.033 (0.046)	0.148 (0.045)
Adjusted R-Squared	0.38	0.39	0.38	0.39	0.70	0.73	0.71	0.73
SIC3 Fixed Effects	X	X	X	X	X	X	X	X

Notes: See Table 3.

App. Table 6: Conditional Estimations of Mfg. Entry Counts

	Unweighted Estimations					Weighted Estimations				
	Existing Industry Structure	Add Labor + I/O Flow	Add Chinitz Metric	Add Tech Flows	Add Facility Expansions	Existing Industry Structure	Add Labor + I/O Flow	Add Chinitz Metric	Add Tech Flows	Add Facility Expansions
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
<i>Dep. Variable is Log Entry Count of New Firms by City-Industry</i>										
Log of Employment in City-Industry	0.120 (0.006)	0.056 (0.006)	0.047 (0.006)	0.037 (0.006)	0.042 (0.005)	0.406 (0.048)	0.020 (0.038)	-0.071 (0.036)	-0.104 (0.035)	-0.008 (0.029)
Labor Market Strength Metric		0.086 (0.010)	0.070 (0.010)	0.022 (0.011)	0.017 (0.009)		0.679 (0.073)	0.395 (0.065)	0.230 (0.070)	0.200 (0.054)
Inputs / Supplier Strength Metric		0.096 (0.011)	0.042 (0.013)	0.026 (0.013)	0.004 (0.011)		0.262 (0.071)	-0.098 (0.073)	-0.174 (0.071)	-0.177 (0.051)
Chinitz Measure of Small Suppliers			0.107 (0.016)	0.107 (0.016)	0.092 (0.014)			0.979 (0.075)	0.954 (0.073)	0.611 (0.049)
Outputs / Customer Strength Metric		0.048 (0.012)	0.040 (0.010)	0.033 (0.010)	0.020 (0.008)		0.046 (0.039)	0.094 (0.036)	0.069 (0.034)	-0.004 (0.021)
Technology Strength Metric				0.122 (0.013)	0.090 (0.012)				0.411 (0.056)	0.266 (0.048)
Log Facility Entry by Multi-Unit Firms					0.336 (0.012)					0.186 (0.008)
Adjusted R-Squared	0.45	0.45	0.46	0.46	0.56	0.76	0.77	0.79	0.79	0.84
City Fixed Effects	X	X	X	X	X	X	X	X	X	X
SIC3 Fixed Effects	X	X	X	X	X	X	X	X	X	X

Notes: See Table 4.

App. Table 7: Conditional Estimations of Mfg. Entry Counts by Entrant Size

	Unweighted Estimations					Weighted Estimations				
	Total	Entering Employment Of				Total	Entering Employment Of			
	Entry	1-5	6-20	21-100	101+	Entry	1-5	6-20	21-100	101+
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
<i>Dep. Variable is Log Entry Count of New Firms by City-Industry</i>										
Log of Employment in City-Industry	0.037 (0.006)	0.007 (0.006)	-0.014 (0.006)	-0.013 (0.008)	-0.006 (0.008)	-0.104 (0.035)	-0.243 (0.040)	-0.439 (0.063)	-0.579 (0.143)	-0.049 (0.069)
Labor Market Strength Metric	0.022 (0.011)	0.018 (0.011)	0.030 (0.013)	0.024 (0.017)	0.019 (0.020)	0.230 (0.070)	0.263 (0.096)	0.407 (0.163)	0.627 (0.369)	0.426 (0.405)
Inputs / Supplier Strength Metric	0.026 (0.013)	-0.002 (0.013)	0.023 (0.014)	0.040 (0.020)	-0.022 (0.027)	-0.174 (0.071)	-0.277 (0.082)	-0.334 (0.139)	-0.607 (0.370)	-1.234 (1.133)
Chinitz Measure of Small Suppliers	0.107 (0.016)	0.108 (0.017)	0.070 (0.013)	0.031 (0.010)	0.009 (0.007)	0.954 (0.073)	1.026 (0.092)	1.320 (0.164)	1.823 (0.351)	0.990 (0.803)
Outputs / Customer Strength Metric	0.033 (0.010)	0.032 (0.010)	0.037 (0.013)	0.052 (0.024)	0.047 (0.047)	0.069 (0.034)	0.073 (0.034)	0.176 (0.069)	0.463 (0.199)	0.465 (0.433)
Technology Strength Metric	0.122 (0.013)	0.120 (0.014)	0.110 (0.017)	0.075 (0.019)	0.021 (0.032)	0.411 (0.056)	0.489 (0.075)	0.668 (0.132)	0.775 (0.269)	0.542 (0.563)
Adjusted R-Squared	0.46	0.38	0.24	0.10	0.00	0.79	0.74	0.62	0.39	0.07
City Fixed Effects	X	X	X	X	X	X	X	X	X	X
SIC3 Fixed Effects	X	X	X	X	X	X	X	X	X	X

Notes: See Table 5.

App. Table 8: Conditional Estimations of Mfg. Entry Counts through Natural Advantages

	Unweighted Estimations			Weighted Estimations		
	Existing Industry Structure	Add Labor + I/O Flow	Add Tech Flows	Existing Industry Structure	Add Labor + I/O Flow	Add Tech Flows
	(1)	(2)	(3)	(4)	(5)	(6)
<i>Dep. Variable is Log Entry Count of New Firms by City-Industry</i>						
Pred. Log of Empl. in City-Industry	0.397 (0.033)	0.391 (0.036)	0.346 (0.037)	1.993 (0.283)	1.070 (0.343)	1.193 (0.339)
Pred. Labor Market Strength Metric		0.125 (0.059)	0.064 (0.060)		1.425 (0.377)	1.514 (0.442)
Pred. Inputs / Supplier Strength Metric		-0.232 (0.061)	-0.226 (0.060)		-0.373 (0.306)	-0.344 (0.318)
Pred. Outputs / Customer Strength		0.033 (0.037)	0.033 (0.036)		0.704 (0.255)	0.727 (0.262)
Pred. Technology Strength Metric			0.184 (0.058)			-0.252 (0.443)
Adjusted R-Squared	0.44	0.44	0.44	0.75	0.76	0.76
City Fixed Effects	X	X	X	X	X	X
SIC3 Fixed Effects	X	X	X	X	X	X

Notes: See Table 6.