Entrepreneurial Migration*

Kevin A. Bryan[†]

Jorge Guzman[‡]

August 30, 2021

Abstract

We use cross-state business registrations to track the geographic movement of startups with high growth potential. In their first five years, 6.6% percent of these startups move across state borders. Though startup births are concentrated geographically, hubs like Silicon Valley and Boston on net *lose* startups to entrepreneurial migration. A revealed preference approach nonparametrically identifies the average utility of cities to migrant founders. University towns and startup hubs have *low* relative utility. This pattern is due neither to vertical sorting nor industrial specialization. The higher-quality startups move to lower-tax, business-friendly cities, while less growth-oriented startups move to low-tax, high-amenity cities.

*We thank Yupeng Liu for outstanding research assistantship on this project, and acknowledge the generous financial support of the Kauffman Foundation through their Uncommon Methods and Metrics initiative. We thank Thomas Astebro, Jeff Furman, Ruben Gaetani, Joshua Gans, Olav Sorenson, Scott Stern, and Rosemarie Ziedonis, as well as seminar participants in Columbia Business School, Harvard Business School, Boston University, the University of Toronto, the Wharton Migration and Organizations Conference, and the UNC Emergence Conference. All errors and omissions are our own.

[†]University of Toronto, Rotman School of Management. kevin.bryan@rotman.utoronto.ca.

[‡]Columbia Business School. jag2367@gsb.columbia.edu.

1 Introduction

Most high-growth startups are born in a very small set of cities (Guzman and Stern, 2020). These traditional “startup hubs” tend to possess fixed resources which are very challenging for other cities to replicate, like major technical universities, keystone firms spinning out founders, or a venture finance ecosystem. However, in addition to *birthing* startups, there is a second option available to cities: *attract* these young and fast-growing firms to move from their initial location. While most startups remain close to where they are born (Dahl and Sorenson, 2012; Michelacci and Silva, 2007; Guzman, 2018), a number of prominent firms, including Microsoft, Facebook, and Slack, moved when young. How common is this entrepreneurial migration? What cities do they prefer? What leads startups who move to choose one city over another?

Policymakers in practice, and canonical models of spatial economics in theory, offer wildly different explanations for firm mobility. Consider the following three cases, each of which is not simply a vignette, but also a datapoint in our empirical analysis:

Tableau Software, a visualization and data analytics company, was founded in 2003 by a group of Stanford researchers. The next year, they moved their company headquarters from Silicon Valley to Seattle, where the company would grow before being acquired for \$15.7 billion by Salesforce. The founders argued it was simply a lifestyle decision. Both wanted to live in Seattle even though “[i]t’s clearly no Silicon Valley in terms of sheer volume of technology companies.”¹

Stason Animal Health was founded in 2011 in the suburbs of Portland, Oregon, as a venture-backed company focused on pharmaceuticals for pets. In 2013, they moved their headquarters to Kansas City, Kansas, attracted by the booming ecosystem for their industry in the ‘KC Animal Health Corridor’. “The culture of Kansas City and concentration of animal health companies here made the selection quite easy. There is no place like the KC Animal Health Corridor for a company looking to serve the animal health industry,” remarked the CEO Diana Wood.²

Vbrick Systems launched a video platform in 2015 following a pivot from a focus on video encoders, which are a hardware product. The new CEO following this business model switch moved the company’s headquarters to Herndon, Virginia

¹<https://xconomy.com/seattle/2008/09/08/tableau-raises-10m-in-second-venture-round-wants-to-be-the-adobe-of-data/>

²<https://www.kansascity.com/news/local/article326075.html>

from Wallingford, Connecticut “to get access to technical talent in the D.C. area.” The firm quickly raised \$20 million to expand sales and marketing of the new product.³

These examples illustrate a variety of reasons for startup mobility, including founder preferences (Tableau), Marshallian agglomeration effects (Stason), and labor availability (Vbrick). They touch on aspects of both firm productivity and founder utility, the potential substitutability of these two dimensions, and their diverging consequences for startup performance. For example, founders prioritizing amenities may move to sunny locations with lower-quality labor, while founders seeking strong ecosystems may move to cities which are colder or rainier as long as the talent is sufficiently strong.

Part of the reason that entrepreneurial migration is so poorly understood is that it is very challenging to track the migration of high-growth startups. They are often too small or young to appear in censuses or other standardized datasets. Even when they can be identified, it is challenging to separate growth-oriented young firms from laundromats and pizza parlors.⁴ We use a technique developed in Guzman and Stern (2015) and Guzman (2018) to cull high-quality startups using the information in their initial state-level business registration. We then take advantage of the uniqueness of names in the Delaware business registry to trace cross-state moves. This method lets us track changes in the headquarters location of startups with high growth potential at birth (“startups”). Our data includes the universe of Delaware-registered startups born between 1988 and 2014 in 36 jurisdictions (35 states plus the District of Columbia, hereafter referred to as “states”) representing roughly 82% of the U.S. population.

We begin in Section 2 by documenting three stylized facts. First, entrepreneurial migration is common; indeed, these firms move across state lines slightly less often than working-age adults. Second, the rich do not get richer: more startups move *out* of major hubs like

³<https://technical.ly/dc/2018/06/20/vbrick-20-million/>

⁴The vast majority of new firms do not intend to grow. Separating migration of potentially high growth firms from, for example, an LLC holding an individual’s investments is particularly important in our setting. See Schoar (2010) and Hurst and Pugsley (2011) for more on this point.

Boston, San Francisco, and Silicon Valley than move *to* these hubs. Third, low-tax cities are relatively attractive to movers versus founders, with important caveats.

In section 3, we formally model the underlying average utility of a city for entrepreneurs directly from revealed preference. Our model of firm location choice modifies a technique employed by Sorkin (2018) in the context of worker revealed preference over companies. The bilateral patterns of movements across cities identifies the average utility to cities of moving firms relative to that which justifies the pattern of startup births. In particular, even when most city pairs have zero moves between them, and even when some bilateral pairs are missing in the data, we can nonetheless use the information in the *network* of moves to recover the average utility of each city. Further, these utilities are recovered analytically and nonparametrically via an application of the Perron-Frobenius Theorem. The estimated rank of city utilities for migrating entrepreneurs does not therefore require any prespecification by the analyst of explanatory covariates. Roughly, the model will suggest that one city is “better” for entrepreneurs than another if it attracts firms from other good cities, and loses few firms.

In section 4, we describe the dataset we draw on in more detail. We track 27 years of headquarter locations of over 400,000 startups with high ex-ante growth potential. Unlike many other studies of young firms, we do not restrict to VC-funded firms, to firms in industries well-covered by censuses such as manufacturing, or to definitions of young firms where many have limited at-birth likelihood of ever growing (Shane, 2009). As our dataset does not depend on future sales or other aspects of performance correlated with location, we can track moves without concerns about post-birth selection.

Section 5 contains our primary empirical results. The recovered vector of utilities shows a striking pattern: as was suggested by our second stylized fact, the utility to movers looks quite unlike the relative ranking of cities by high-quality startup *births*. The highest ranking large cities for movers are Dallas, Phoenix, Austin, and Charlotte. University towns, poorly educated small cities, and startup hubs like Boston and the Bay Area all have lower than

average utility for these firms. The pattern of city moves is also not consistent with a model where firms predominantly move for idiosyncratic reasons with all cities possessing identical common utility: there are in fact cities that are disproportionately attractive or unattractive for migrating entrepreneurs.⁵ Furthermore, when we compare the estimated utilities of migrant Delaware corporations to LLCs, who are less growth oriented on average, we observe striking differences amongst the two. This suggests that estimated mover utility in each city, and hence the factors which drive differential migration across them, is sensitive to the growth orientation of the firms being examined.

To expand on this point, startups are unusual in that they face a fundamental tension between the preferences of the founders, and the profitability of the firm. One might therefore imagine that startups, when choosing where to locate, hold an intermediate position between individual workers selecting a new city and established firms selecting a new plant location. The individual worker selects a location on the basis of wages as well as non-pecuniary benefits of a city like nice weather or its “Bohemian” nature. Established firms select plant location based on the local labor pool, tax advantages, and other formal incentives. By examining the estimated utility from revealed preference of cities to entrepreneurs, both high- and low-quality, we can better understand the factors driving their spatial choices.

In particular, we consider how consistent our estimated mover utility is with a number of common theories of firm location. The *Marshallian explanation* is that firms agglomerate due to shared resources, either at the industry level (Marshall 1920) or the business services level (e.g., Ota and Fujita (1993)). The *Jacobs explanation* is that, instead, industrial diversity in a region improves firm performance by providing a fertile ecosystem of new ideas (Jacobs,

⁵As the theoretical section clarifies, we use “idiosyncratic” to mean factors determining migration which depend on preferences of individual firms rather than a common component of utility shared by all firms. This common component may, and does, contain factors beyond pure profit potential. Other authors studying location choice, such as Dahl and Sorenson (2009), define idiosyncratic as all non-pecuniary motives, including things like a preference for sunny weather which may be widely shared by founders.

1969).⁶ These two explanations combine in the “*nursery cities*” model of Duranton and Puga (2001), where diversity of ideas is useful for young firms, but Marshallian agglomeration is useful once those firms have figured out a product and business model. The *ecosystems explanation* is that “startup ecosystems” provide Marshallian spillovers to other startups through specialized labor, capital, and culture.⁷ The *amenities explanation* is that because founders and early employees are mobile, they will move to high amenity cities and avoid high cost of living cities unless they receive a large compensating differential in terms of performance (Roback, 1982; Albouy, 2016; Guzman, 2018). The *Bohemia explanation* builds on the work of Florida (2002) to posit that startups are attracted to particular amenity mix found in “bohemian” cities. The *tax policy* explanation argues that startups, like established firms and inventors, migrate to reduce either their tax burden or the tax burden of their founder-owners (Moretti and Wilson, 2017; Fajgelbaum et al., 2019).

Empirically, we find that startups founded as corporations derive little relative utility from cities with high education, including “college towns”, or from cities with a diverse industrial mix when they choose to move. They prefer low-tax cities, and if anything avoid high-amenity locations. In contrast, for LLCs, we see a meaningful positive response to amenities, especially low housing costs, sunshine, and composite quality of life measures, to low tax rates, and to the factors which led many LLCs per capita to be born in the destination city. The tax response for LLCs is stronger at median incomes, whereas the tax response of corporations is strongest at the right tail.

We should note that while our *ranking* of cities by their relative utility to movers is nonparametric, care should be taken in interpreting these city-level factors as causal. Further, these results should not be seen as “tests” of urban agglomeration theories outside the context

⁶See Beaudry and Schiffauerova (2009) for a review of the strengths and weaknesses of these two models in explaining empirical firm location patterns. They particularly note the difficulty in separating diversity from co-agglomeration when industry markets are relatively aggregated. That is, while diversity within the mining sector may provide Jacobs effects, they look like Marshallian effects if industry can only be coded as “mining”, “agriculture”, or “industry”.

⁷This bears some resemblance to the Porter (1990) argument that competitive cities drive higher performing firms, at least partially through the positive spillovers of better labor, management, and firm-level innovation.

of young firms. Rather, we are interested in which theories are sufficiently first-order to explain the particular empirical pattern of startup births and moves.

1.1 Relation to Existing Literature

This paper builds on the literature in both urban economics and entrepreneurship.

First, it provides important empirical evidence on migration and the value of locations for migrant startups. Though there is a large literature on the birth and evolution of entrepreneurial clusters (Saxenian, 1994; Michelacci and Silva, 2007; Delgado et al., 2010; Kerr and Robert-Nicoud, 2020; Chatterji et al., 2014; Glaeser and Kerr, 2009), the vast majority focuses on the differences in local characteristics that lead to different levels of firm formation. Only recently have a small number of papers begun to consider migration (e.g. Dahl and Sorenson, 2012; Guzman, 2018; Conti and Guzman, 2019), but this work has remained purely focused on identifying the impact of moving on individual startup performance. Relative to this prior work, our contributions include the first systematic measurement and benchmarking of startup migration rates in the United States, and a new way to use the network of moves to understand the underlying value of all destinations. We report that startup migration is indeed relatively common. Our approach allows us to characterize the value of each city using relatively weak assumptions, and report the most valuable cities throughout the United States.

Second, we also use these results to understand which urban characteristics correlate with startup mover utility. This allows us to shed light on how various theories of agglomeration might explain the migration choices of high growth startups. Because startups are a key driver of regional and national economic growth (Glaeser et al., 2015; Haltiwanger et al., 2013), their desired urban characteristics are critical to understanding urban productivity.

At a policy level, these results emphasize the presence of geographic misallocation of productive activity in the United States (Hsieh and Moretti, 2019) and the important role of migration in mediating it. Destinations with high livability attract entrepreneurial migrants,

but are not particularly attractive to the high-growth set of Delaware-registered corporations which includes virtually all venture-backed startups. As well, while local idea diversity, bohemian culture, a highly educated population, and the presence of universities play a central role in producing a large supply of new startups, they are not particularly attractive for migrants. Consistent with evidence on innovator location choices (Moretti and Wilson, 2017; Akcigit et al., 2016), personal tax rates appear particularly important.

To be clear, our results are about what factors are *relatively* important for migrants compared to the factors that cause cities to vary in how many entrepreneurs they create. We are agnostic when it comes to the importance of any of these factors on shifting the nature of human capital in a region, or the choice residents make between paid labor and entrepreneurship. Nonetheless, the primary policy interest in startup migration involves attracting firms to cities which otherwise have trouble generating high-growth entrepreneurship locally, and the primary managerial question involves understanding how startups operate geographically after they exist: in both cases, the relative utility we estimate is the most appropriate theoretical construct.

2 Three Facts about Entrepreneurial Migration

Before explaining our data in detail, we begin by documenting three facts on the migration of startups across the United States. These empirical regularities will frame both the basic facts of the phenomenon of study, and highlight some key differences between the net distribution of entrepreneurship and the destination choices of startups that choose to move. As we will discuss in greater depth in Section 4, “startup” here refers to a firm registered in Delaware at birth as either an LLC or a corporation. Delaware registration at birth is highly indicative of growth intention (see also Andrews et al. (2020) and references therein).

Fact 1. *6.6% of startups move across state-lines to different cities (metropolitan areas) within five years of founding.*

This estimate is our baseline estimated move rate of startup migration for those born between 1988 and 2014, after making a few assumptions to account for the fact that we only observe migrations between 36 states. In particular, for each source-destination state pair in our data, we run a regression on the number of movers using 7 controls: the source and destination populations in 2010, the ratio of the two, the source and destination population squared, and the source and destination population growth since 1990. We then use the predicted value of this regression to estimate the number of movers leaving each of our data states to destinations we do not observe. Our estimated startup migration rate is slightly lower than the 5-year interstate migration rate for individuals in the US, which was 8.0% in 2005 and has decreased significantly since (Frey, 2017).

Fact 2. *Many important startup hubs lose more startups to migration than they gain.*

Figure 1 plots the net migration ratio—the number of arrivals divided by the number of departures—against two measures of the strength of the local entrepreneurial ecosystem. Panel A compares with the total number of startups founded in each city. We observe the largest cities in our data at the top of this measure (Los Angeles, New York, Chicago, etc). There is little relationship between the total number of high growth startups in the city and the net migration rate to the city. The R^2 of the fitted line is 2%. Panel B plots instead the intensity of the entrepreneurial ecosystem by dividing the number of startups by the city size. Consistent with most estimates of entrepreneurial ecosystems, Silicon Valley ranks at the top, followed by San Francisco, Boston, San Diego, and Austin. Yet, we see once again little relationship between net migration rates and a city’s local entrepreneurial intensity.

Figures 2 and 3 show this graphically, plotting startups births and net startup moves per capita in large CBSAs. In both pictures, high per capita figures are represented by red and low per capita figures by blue. Again, the cities with the most startups per capita are what one would expect: San Jose, San Francisco, Boston, New York City, Austin, Los Angeles, and so on. However, many of these locations lose more startups in their first five years, including San Jose and San Francisco. Yet many sunbelt cities, Charlotte, Seattle and

Minneapolis attract many more startups than they lose.

We will show in the following section that these patterns can be interpreted precisely, under fairly nonrestrictive assumptions, by extracting the average utility of each city for all potential movers *relative* to the spatial distribution of utilities that would justify the initial distribution of startup locations. These revealed preference-based relative utilities imply the following:

Fact 3. *A strong predictor of which cities gain startups is their personal income tax rate.*

Figure 4 shows that a lower personal income tax rate, measured here at the 95th percentile using the computed data in Moretti and Wilson (2017), is highly predictive of net moves. This pattern again will hold up when we compare tax rates to a precise measure of the relative utility of a city to a moving startup. As we shall see, this comparative static represents a key takeaway of our paper: entrepreneurs’ location choices relate to both amenities that are consumable by utility maximizing individuals and the productivity of their firms. It suggests a view of entrepreneurs as rational individuals with a utility function that only partially depends on the success of their firms. Though the distribution of new startups *founded* by city is highly skewed, we do not find evidence that cities with high propensity to generate startups have equal ability to attract them. Indeed, in the case of university towns, the opposite is true: though these cities spin many startups out of their universities, they do not attract firms from the outside to join that ecosystem. We unpack the implications of these results in Section 5.

3 A Revealed Preference Model of Startup Migration

How should we investigate whether startups prefer to move to city A or city B, and why? A common method for questions of this type is to gather a set of city-level covariates X_i , the distance between any city pair D_{ij} , and the city population P_i , then estimate a *hedonic gravity model* (e.g., Guimaraes et al. (2003), Schmidheiny and Brulhart (2011)) like

$$y_{ij} = \frac{P_i \times P_j}{D_{ij}} e^{X\beta} e^{\beta_0}$$

This is a gravity model, because, like Newtonian gravity, flows of startups y_{ij} between two cities are higher when those cities’ populations P are higher, and when their distance from each other D_{ij} is lower. This is a hedonic model because, aside from the gravity features, it attempts to explain flows using origin- and destination-level covariates which might attract or push away firms. City-level covariates might include education levels, the existence of startup communities born in the area, whether the city is a “university town”, the education level of the city, natural features such as sunlight, heat and cold, per capita incomes, housing prices as a proxy for general cost of living, the industrial concentration, and so on.

The main problem with fitting a model of this type is that it is not at all clear which city-level observables explain the variation in firm move behavior. It is difficult a priori to construct a complete list of relevant covariates, which means that any estimate like the one above will suffer both from omitted variable *bias* in the statistical sense, but also omitted variables in the explanatory sense. Indeed, it may be the case that important features of what startups care about are data completely unavailable to the econometrician. Further, it is not clear how to map from shifts in flows to shifts in founder preferences, limiting the ability to make welfare claims or investigate counterfactuals.

There is an alternative, however, if we return to the first principles of revealed preference. In particular, we will assume that the spatial distribution of startups at birth is driven by idiosyncratic firm-level factors plus common city-level fixed effects. We then assume that when startups are young, with a fixed probability they consider moving. If they do so, they receive another set of utility draws for all cities including their home city, and move to wherever they get the highest draw.

This model induces a network structure, where moves between cities A and B, and B and C, are informative about the relative average utility of A versus C. Assume we only observe data on, for every pair of cities, the number of bilateral moves in each direction.

The *full network* of all moves helps back out a revealed preference ranking among *all* cities even though we only observe a small number of possible pairwise comparisons. We will then have a rank-order of utility constructed without any a priori assumptions about what features drive startup choice. This utility ranking can then be brought either to second-stage explanatory regressions, or to direct rank-order comparisons with alternative MSA- or state-level rankings of cities along some other dimension (its bohemian nature, its business climate, its natural amenities, and so on).

The technique we use here is a modification of one developed in Sorkin (2018) for the purpose of understanding the non-wage component of jobs from different employers, given data on wages and firm-to-firm voluntary transitions. The model requires two assumptions, one about how startups are born, and one about how they decide to move.

Assumption 1. *Assume there are N potential entrepreneurs in society. Each potential entrepreneur i receives a utility draw of $B_j + \mu_{ij}$ from beginning a new startup in each city $j \in J$ and a draw $B_0 + \mu_{0j}$ from a null option of not forming a startup. Let each μ be a draw from a mean-zero i.i.d. extreme value type 1 distribution with scale 1.*

The first assumption says that there is some otherwise unmodeled rationale for the observed pattern of startup births. The literature on firm formation often assumes that differential startup rates across cities depend on factors like the number of existing firms a startup could spin out of, the number of potential founders living in the city, and so on (e.g., Buenstorf and Klepper (2009), Saxenian (1994)). For instance, if higher populations birth more startups, then ceteris paribus B will be higher in cities with more people.

Using standard results from discrete choice, the expected number of firms born in city j is

$$N \frac{e^{B_j}}{e^{B_0} + \sum_{k \in J} e^{B_k}}$$

Assumption 2. *After they are born, let a fraction δ of all startups receive another utility draw from each city, equal to $V_j + \epsilon_{ij}$. As before, ϵ is a draw from a mean-zero i.i.d. extreme*

value type 1 distribution with scale 1.

That is, we assume a fraction of all startups consider moving after birth. If they do, they get a new utility draw from all cities, and move to the city with the highest draw (or else stay in their birth city if that is maximal). We will call V_j the “common utility” component of a city’s utility to a startup, and ϵ_{ij} the idiosyncratic component.

The fact that the error term is uncorrelated across cities is an important assumption: it rules out that, for instance, cities with similar industrial bases have correlated utility for a given firm beyond that which drives common value for all firms. That is, if firm moves are simple industrial sorting, this model is inappropriate: cities in a given industry give firms correlated utility, and this correlated utility is not common utility because it applies only to firms in that specific industry.⁸ However, if the industrial *diversity* of a city, or the level of industrial specialization, or the amenity value of cities is what drives firm moves, and firms merely differ in the importance they place idiosyncratically on those features, the assumption holds.

We now show how to analytically extract the vector of *relative utilities* $\bar{V} = V - B$. We call these relative utilities because they are the attractiveness of a city for a startup relative to that which would keep the spatial distribution of startups constant after accounting for startup mobility. If $\bar{V}_j$ is positive, startups get more common utility from a city than that which would rationalize the at-birth spatial distribution. That is, we are interested in identifying which cities have factors more conducive to attracting movers than to birthing startups.

Importantly, we are able to identify $\bar{V}$ even if we do not know the size of the potential entrepreneur set N , the average utility of abstaining from entrepreneurship B_0 , or the probability startups consider moving δ . All that we will require in the data is that for every city pair $\{j, k\}$, we observe the total number of bilateral moves between that pair. Further, the

⁸Note that it is possible to perform the algorithm described in this section industry-by-industry to examine the extent of heterogeneity in common utility values. As we do not observe industry, and can only guess it based on firm name with nontrivial error, we do not further probe this assumption.

model identifies the relative utility of cities even when there are no bilateral moves between some city pairs, as long as the network of moves between all cities is strongly connected. The fundamental idea is that even if we observe no direct moves from Shreveport to Spokane, or vice versa, the former is more attractive if we observe firms from Spokane moving to Biloxi and firms from Biloxi moving to Shreveport. Finally, the model is identified even if we do not observe bilateral moves for some subset of cities, such as within-state moves or moves to or from the 15 missing states in our data.

Let us now derive $\bar{V}$. The number of firms who *move* from city j to city k is

$$M_{jk} = N \frac{e^{B_j}}{e^{B_0} + \sum_l e^{B_l}} \delta \frac{e^{V_k}}{\sum_l e^{V_l}}$$

That is, the number of movers from j to k is equal to the number of firms born in j times the probability a given firm considers moving times the probability it gets its highest post-birth utility draw in k . We therefore have that

$$\frac{M_{kj}}{M_{jk}} = \frac{e^{V_j} e^{B_k}}{e^{V_k} e^{B_j}} = \frac{e^{V_j - B_j}}{e^{V_k - B_k}} = \frac{e^{\bar{V}_j}}{e^{\bar{V}_k}}$$

Letting $W_j = e^{\bar{V}_j}$, we have

$$\frac{M_{kj}}{M_{jk}} = \frac{W_j}{W_k}$$

That is, in any bilateral comparison, the city that attracts the most net moves is expected to have higher utility. Again, we often have no, or very few, moves between any given pair of cities. However, expanding from two cities to all cities, we can sum over j on both sides to get

$$\sum_j M_{kj} W_k = \sum_j M_{jk} W_j$$

and hence

$$\frac{\sum_j M_{jk} W_j}{\sum_j M_{kj}} = W_k$$

The denominator is the number of firms born in k that leave, and the numerator is the number that come, weighted by the “utility” of where they come from. If firms from good places come, it’s a better signal of quality than if firms from bad places come. If many people come and few leave, it’s a better signal of quality than if many come and many leave.

This is simply one linear restriction for each firm. As in Sorkin (2018), the model is overidentified since there are also the pairwise comparisons above (most of which are very noisy and many of which are bilateral zeros). Note that when we compare, for example, New Orleans to Seattle, all the firms that choose between New Orleans and city B, or Seattle and city C, will also give information about the value of New Orleans and Seattle since they form part of a “network” of revealed preference of the relative common value portion of city utility $\bar{V}$. This is particularly useful for identifying the relative utility of cities with few total movers, often because they have a small population. For instance, if a small city attracts only one firm, but that firm comes from a city that otherwise loses very few companies, the model puts more weight on the small city being an attractive place rather than one that got idiosyncratically lucky. If the firm it attracts comes from a city that otherwise is fairly unattractive, that one incoming firm may on the other hand be a fairly uninformative signal about how firms on average view the receiving city.

Let us now show how to extract the relative utility vector W from that equality. In matrix form, those linear restrictions can be written $SW = W$ where W is the vector of city common value relative utilities and S is a matrix where $S_{jk} = \frac{M_{jk}}{\sum_n M_{jn}}$. Left to prove is that there exists a matrix W satisfying that equation.

If M is strongly connected, meaning that there is a directed path from every city to every other city in the adjacency matrix based on M , then the Perron-Frobenius theorem applies. Perron-Frobenius says that for irreducible non-negative matrices (e.g., strongly connected adjacency matrices), there is a unique largest eigenvalue whose eigenvector is strictly positive.

That is, there exists a unique solution to

$$SW = \lambda W$$

where λ is the largest eigenvalue and W is the corresponding eigenvector. It is well known that when you apply Perron-Frobenius to a probability transition matrix, then the biggest eigenvalue is equal to 1, and hence we are done: we have solved for W just by finding the corresponding eigenvector to the first eigenvalue.⁹ Given that the eigenvector represents relative values of W , we can convert into city relative common utilities by $\bar{V}_j = \ln(W_j)$, using the definition of W .¹⁰

This method extracts utility nonparametrically, hence in a manner well-suited for the present problem where we do not have good priors for what parametric factors mobile startups care about. In addition, the method is particularly well-suited to data like ours where only a subset of move data is available, but for which the bidirectional flows are always available whenever the unidirectional flow is. The reason is that this estimated common utility of a city under Assumptions 1 and 2 is independent of N , the number of potential entrepreneurs, δ , the fraction of firms that consider moving, and the fraction of firms born in a given city who do not move. Note that since we do not observe firm deaths, within-state moves, or moves to the 15 missing states, we do not actually know what fraction of firms in a given city do not move, so it is essential that our empirical technique does not rely on

⁹In this discrete choice setting, there is one more fairly simple step to prove the biggest eigenvalue is 1 (see Sorkin (2018), Appendix E).

¹⁰Rather than extracting relative common utilities using Assumption 1, one can instead assume that the probability a firm gets a utility draw V_j at all is directly proportional to the at-birth startup population of city j . That is, let $M_{jk} = NPr(V_k > V_j)f_jg_k$, where f_j is the proportion of firms born in j and g_k is the probability any given firm gets a utility draw from city k . If $\frac{f_j}{g_j}$ is a constant for all cities j , these terms and N drop out of the ratio $\frac{M_{jk}}{M_{kj}}$, and the utility V can be extracted following the procedure above. We prefer the interpretation of the model in terms of utility relative to that which rationalizes at-birth firm location, and will rely on that interpretation in Section 5 when discussing implications for agglomeration theory. Nonetheless, we wish to be clear that there exist alternative, analytically-tractable sets of assumptions to infer startup utility from move data. Due to the absence of prices, the method of directly identifying consideration sets simultaneously with preferences as in Abaluck and Adams (2016) does not apply in our setting.

knowing that figure.

Why are our utility estimates independent of the fraction of firms in a given city who do not move? Mathematically, the estimated utility vector $\bar{V}$ is based on an eigenvector whose value is constant regardless of the number of non-movers M_{jj} .¹¹ The intuition here is that since we are estimating utility of a city to movers relative to the utility which would rationalize the initial distribution of firms, and since idiosyncratic draws are uncorrelated across cities for a given firm, relative bilateral flows wholly identify utility asymptotically: a city j with positive net flows from a city k is higher utility with certainty as the sample grows large. The overidentifying assumptions we get from having a larger sample of MSAs helps identify relative utility between city pairs even when they have a small number of bilateral flows but a large number of flows to other cities in the network.

One caveat is that the model requires a strongly connected matrix of moves. We restrict analysis to MSAs with at least 2 firms moving in or moving out, and directly check that the matrix of moves is invertible under that assumption.¹² This constraint binds particularly for LLCs, which have less mobility than corporations. In all tables, we denote by “N/A” the utility of cities which are dropped because of limited LLC data.

4 Measuring Entrepreneurial Migration

Bringing this model to data requires consistent measures of startup bilateral moves for every pair of cities being considered. Measuring this movement of entrepreneurs and their startups across locations is difficult for several reasons. First, in contrast to established companies with set offices and working locations, young entrepreneurs can work at a variety of locations without any of them being common enough to be considered the firm’s place of business. For example, an entrepreneur can spend some time at a coffee shop, some time

¹¹Since $\bar{V}$ is completely determined by linear equations of the form $\sum_j M_{kj}W_k = \sum_j M_{jk}W_j$, the diagonal element M_{jj} appears as $M_{jj}W_j$ on both sides and hence cancels out.

¹²Firms with at least two moves out and no moves in are assigned the utility of the lowest city that is otherwise estimated by the procedure above.

in a co-working space, and some time working while traveling. In this case, the location of the firm itself is unclear. Second, even for those firms that establish a location, observing this location choice is challenging because young startups often leave little observable trace of where they are in formal databases. Finally, third, even if we are able to observe the startups, there is the perpetual concern of startup quality (e.g., Guzman and Stern (2020)): startups are heterogeneous in their underlying potential, most are not growth-oriented, and an approach to studying growth startups independent of their location requires accounting for that orientation at birth.

4.1 Measuring migration through public records

To avoid those issues, we take advantage of the business registration records created when firms are founded. Business registrations are public records created endogenously when a firm is registered as a corporation, partnership, or limited liability company, with the Secretary of State (or Secretary of the Commonwealth) of any U.S. state (or commonwealth).

Taking advantage of the unique institutional setting in the United States, where states are individual jurisdictions and require firms to register in each state in which it does business, we can use the registration process of firms across states to observe the cross-state migrations of startups. Specifically, business registration records require startups to include up to four different addresses of record: the local office in the state of registration, the principal office of the business (i.e., the headquarters), the office of the registered agent (i.e., the lawyer), and the address of the registered directors. While not all addresses are included in all cases, we identify 36 states in which we can identify the principal office of business independently from the local office of business.

We examine startups *that change the principal office of business* in these state datasets to identify headquarter migrations across locations. We identify as a migration any observation for which we can establish three facts: (i) the company with the same name and legal

jurisdiction¹³ has registered in two different states; (ii) the company has changed the location of principal office from an address located in the origin state to an address located at the destination state; and, (iii), there is a gap of at least three months in the time between when the company was registered in the original (founding) state and the destination.

When do we consider a firm to have moved? If a firm is registered in state A, appears in the business registry of state B at least three months after that original registration date, and has a principal business address (or equivalent) at an address in state B, we consider the firm to have moved from state A to state B. We consider the move date to be the date at which this firm first registered to do business in state B. Full details of this process, including how it differs from commercial business registries which generally do not identify startups as young as the ones in our dataset, are given in the Online Appendix.

While this approach can be in principle applied to all companies in the business registries, we focus on a smaller sample of companies that show two markers of growth-orientation at founding: registering as a corporation or LLC, and registering under Delaware jurisdiction rather than with their home state. In the process of choosing a jurisdiction for their company, growth oriented founders benefit from registering the firm in Delaware for several reasons. The Delaware General Corporate Law provides a long canon of decisions that are useful in assessing the predictability of complex contracts. The state has an advanced institutional setup to deal with corporate arbitration including its highly reputed Court of the Chancery. The decisions and legal framework of Delaware are generally regarded as pro-business. These benefits are more useful for startups that hope to grow, especially if they plan to interact with venture capitalists.¹⁴

However, being in the Delaware jurisdiction also holds extra costs and requires two registrations (one in Delaware and one in the state of operation), imposing costs that a business

¹³Note that jurisdiction is not the same as the location of business. All companies have a single state jurisdiction, in which they operate as a local company, while they operate as a *foreign* (to the state) company in other locations.

¹⁴In fact, venture capitalists most often require that portfolio companies are in Delaware because their contracts are specifically written for Delaware corporate law.

that expects to be small is likely to deem unnecessary. This creates a natural separating equilibrium, with mostly growth-oriented companies choosing to register in Delaware. Accordingly, while Delaware companies represent only about 4% of all firms, they account for 50% of all publicly listed firms, and over 60% of all VC financing (see Catalini et al., 2019). Delaware firms are also 23 times more likely to achieve an IPO or be acquired than non-Delaware firms Guzman and Stern (2020). Our approach therefore allows us to track the migration of company headquarters (as stated by the entrepreneurs) across states up to the specific destination MSA for all Delaware registered corporations and LLCs—a set with significant growth orientation.

In spite of its potential, this approach does bring some limitations. First, an important limitation of our data is that, due to the use of state registries, we are not able to observe migrations of headquarters across MSAs within the same state. While this could lead to a different migration *rate* for larger or smaller states, our empirical approach identifies relative mover utility using only bilateral moves for each city pair (recall that estimated utilities are independent of δ and M_{jj} , the fraction of firms who consider moving and the average utility of the birth city). Second, our migrations only track the change of legal headquarters, but the way in which companies interact with locations can often be much more nuanced. Companies expand as multi-establishment firms, or work in distributed teams that can include many locations. In this regard, we believe that while we are identifying an important aspect of startup location choice, it is not the only one. Future datasets can improve upon ours, further shedding light on this question. Finally, migration of established startups is only one of the broader set of relocation actions that can happen in entrepreneurship. For example, many individuals might move to locations amenable for startups before becoming entrepreneurs. These relocations will be unobserved in our data. While this is certainly a limitation for the goal of observing all economic migration, we believe it positions the contribution of our data squarely and more clearly on *actual* entrepreneurship at the time it is happening, rather than eventual entrepreneurship.

It is also important to clarify what counts as a startup in our reckoning. A startup is a formal business entity that begins operating for the first time. Mergers that generate a new corporate entity are therefore startups, as are spinouts. In general, it is not obvious whether these types of entities should or should not be called startups, and it is difficult to identify which business registries are spinouts, so we use a conservative definition of startups which includes them. For instance, in 1996, Lucent Technologies was spun out of AT&T, including the famous Bell Labs division. This company was new, and was independent, though it was not “small” in the sense of many startups. As we are interested in growth-oriented startups, we further drop all companies with “Holding” in their name, and all companies with “II”, “III”, “IV”, etc., at the end of their legal name. In our experience, these tend to be financial or real estate holding companies rather than de novo startups.

Though business registries appear on face to be a promising data source for investigating startup behavior, constructing data tracking state-by-state flows from them is not a simple task. Many states do not make full registration data freely available. Records in some states have frequent errors. The firm headquarters location in some cases only lists a lawyer’s address, in which case we rely on alternative measures, such as the MSA address of a majority of corporate directors, to identify the firm’s metro area. We restrict full details of our matching process to the Online Appendix. However, as noted, our empirical method only requires that if we can observe moves from city A to city B, we can also observe those from B to A. This allows us to simply drop the small number of states whose data practices make it particularly burdensome to observe headquarter locations.¹⁵

We secured the business registration records of all companies under Delaware jurisdiction registered in 36 U.S. states through the Startup Cartography Project (Andrews et al., 2020). The Startup Cartography Project is a project measuring the founding (and registration) of

¹⁵For a number of firms, the origin state registration has a headquarters address which is updated to the new state after the move. Because we know the date the firm was originally registered, we can nonetheless identify the *state* it was born in. In these cases, we assign birth MSAs probabilistically: if a firm moving to Dallas is known to be born in Massachusetts but the MSA is unknown, and 80% of known births are in Boston, we assign .8 firm moves from Boston to Dallas. Moves are given in rounded numbers in all tables. Full details of this algorithm are available in the Online Appendix.

all companies in the United States for 49 US states and Washington DC, from 1988 to 2014.¹⁶ From this data, we attempt to extract the local and principal address of office for each firm that also has a Delaware jurisdiction. To acquire the office location, we relied on the formally identified location of firms if possible, and on the location of the majority of firm directors if not. Using this approach, we identified 15 states in which we did not think we were able to adequately separate the local address from the headquarters, which are excluded from the current version of the paper. The states for which move data can be calculated represent 82% of the US population and 86% of the 50 biggest metropolitan areas by population. See Figure 5 for a visual display of the included states.¹⁷

4.2 Other firm and city measures

We complement move information with external data on startups and cities they are considering.

Measures of firm outcomes. While the bulk of our analysis is not based on firm outcomes, we still measure two firm outcomes for the firms in our data, which are used to implement the entrepreneurial quality approach in (Guzman and Stern, 2020). Our first measure, *IPO* is a binary variable equal to 1 if the firm achieves an IPO in the NYSE or the NASDAQ as reported by the SDC Global New Issues database. Our second measure, *Acquisition*, is a binary variable equal to 1 if the firm is reported as being fully acquired in the SDC Platinum Mergers and Acquisitions database.

Measures of Amenities. We construct four measures of amenities that are non-varying across time. *Cooling Degree Days* is the number of days above 65 degrees Fahrenheit weighted by the number of degrees above that point each day, *Heating Degree Days* is the analogue for weighted days below 65 degrees Fahrenheit, *Sunshine Percentage* is the average percentage of

¹⁶Firms that are local to Delaware are excluded from the data due to the unique nature of this state corporate law.

¹⁷Ongoing work is underway to try to recover firm location information in the remaining jurisdictions. The only large missing states are Pennsylvania, Wisconsin, Maryland, and Michigan. We also omit the Trenton, NJ and Augusta, ME firms due to idiosyncrasies in how states record firms in these capital cities.

days that have sunshine, and *Distance from the Water (Inverse)* is the log inverse distance of the city to an ocean or a Great Lake. We also include *Latitude*, to account for any remaining elements of distance to Canada or Mexico. Finally, we include a composite *Quality of Life* index developed by Albouy (2016), based on the revealed preference of individual workers.

Measures of Income Taxes. We include the year-specific measures of the income tax rate in each state derived by Moretti and Wilson (2017). Our four variables on the *Average Income Tax Rate* represent the effective total tax rate (federal and state) for an individual at the 50th, 95th, and 99th percentiles of incomes.

Other Regional Measures from the US Census. Using the US Decennial Census, we also estimate eight measures related to the underlying attractiveness of the entrepreneurial ecosystem for both startups and individuals according to theory. *Bohemia* is an index that replicates the Florida (2002) Bohemian index using the standardized location quotient of the geographic distribution of workers in a few specific ‘bohemian’ occupations. *Industry Concentration* is the Herfindahl index of employment shares across each four-digit industry in a city. *Average Home Value*, *Average Income*, and *Average Wage* are the average value of each of those variables for each city in those observations reported in our census sample. *Share College Educated* and *Share University Students* measure the fraction of population with a university degree or higher, and the fraction of population who are postsecondary students, in each city. We define a *College Town* as one where more than 10 percent of the population are students. *Population* is the population of each city as estimated in decennial censuses US Census. For each value, we take the average of 1990, 2000, and 2010 census estimates.¹⁸

Measures of Venture Capital Financing and Innovation. To account for the underlying financial economy of the region, we create a mapping between CB Insights and our CBSAs, and estimate *Total VC Raised* (in millions) and *The Number of Startups Financed*. We restrict to companies that have raised at a valuation of at least 1 million dollars, to fundraising

¹⁸We construct these estimates on the basis of CBSA county definitions as of 2013, so that the geographic region is held constant across the census years.

between 1988 and 2014, and to fundraising in traditional venture capital equity rounds, seed rounds, or convertible debt.

Patenting. Finally, we use the reported public tables from the USPTO measuring the number of patents granted per year for each CBSA, and include *Patents Granted* as the observed value by the USPTO in each year between 2000 and 2015.

4.3 Summary Statistics

Table 1 presents the summary statistics of all the Delaware-registered corporations in our data. Our dataset includes 181,663 corporations. Of this sample, 0.5% of them have had an IPO and 2.6% of them have been acquired. Highlighting the growth orientation of our Delaware-registered corporations, these figures are more than thirty times higher than the average probability of positive growth outcomes of all new firms, estimated in Guzman and Stern (2020) at 0.07%. Turning to founding characteristics, 5.8% of the companies have a patent at or around founding and 2.4% of the companies have a trademark. Finally, in moving rates, 3.3% of our companies move to a new state within 2 years, 5.6% within 5 years.¹⁹ We also track 237,307 Delaware-registered LLCs, who are much less likely to be acquired (.4%) or to move (2.8% within five years).

We aggregate this data into information on migration flows on two dimensions: state and Metropolitan Statistical Area (MSA), using the 2013 U.S. Census CBSA definitions. The resulting dataset is a matrix containing the number of movers from each source location to each destination location. In any given year, the modal MSA receives 0 high growth startups, and the median MSA receives 1.

Table 2 presents the summary statistics of the state and MSA level flows data. Panel A describes the state to state flows. There are 1,260 possible source-destination state pairs (36 home states moving to 35 other states), with an average number of movers between any unidirectional dyad of 3.7 corporations and 4.2 LLCs. Even within our 27 year period, 47%

¹⁹With comparison to the higher estimate in Section 2, this is raw rather than imputed data. Accounting for estimated moves to states we do not observe, our estimated migration rate for corporations is 6.6%.

of dyads do not have a single move between them.

Panel B is the summary statistics for the MSA to MSA flows. The sparseness of the data is dramatic in this case. Out of a total of 34,040 MSA unidirectional source-destination pairs, only 5.2% (1,749 pairs) have any movers at all. The average number of movers conditional on having at least one move between MSA pairs is 3.8. This sparseness highlights the value of our empirical method. Because the migration graph is so sparse, the key challenge is how to get information within this sparse data to learn more about the value of each city within the totality of cities.

Figure 6 shows the distribution of migration rates for startups across their age profiles. We observe a monotonic reduction in the age probability of moving, decreasing steeply initially and then tapering off. Startups have a 2.1% probability of moving in their first year (age 0), followed by 1.2% probability in the second year and 0.9% in the third. By age 5, this probability has reduced to 0.4% and, by age 10, to 0.2%.

Finally, we document the distribution of migration over our time period in Figure 7. To do so, we plot in the top-left panel the five-year migration rate for each yearly cohort of companies born up to 2009. Two patterns emerge. First, there is a secular decline in the migration rates of startups over time going from 6.9% in 1988 to 5.1% in 2010—a 26% drop in magnitude. The fitted line trend is $-.0008$ and we reject the null that the coefficient is zero (i.e., that there is no decline) at the 1% level using robust standard errors. Second, there is a level of pro-cyclicality around this trend. There are large drops in the migration rate during the years of recession, 1991, 2001, and 2007, and migration rates are relatively higher during the economic boom years. This pattern mirrors a documented secular decline in the inter-state migration rate amongst individuals, as well as other secular drops on business activity more generally (Decker et al., 2014). The top-right, bottom-left and bottom-right panels show this decline in migration holds even if we only look at corporations, or the “highest ex-ante quality” corporations who hold a patent or trademark at founding.

5 Empirical Results

Table 4 shows our primary result. Though our empirical method estimates mover utility for all US MSAs with at least four high-growth startups moving in or out during our 27 year sample, for readability we restrict here to MSAs with a population over 1 million (the listing of all cities by utility can be found in the Appendix). Consider first the raw number of high-growth startups who move into or out of each city. We observe significant variation in the number of movers across states, the relative ratio of inflows to outflows, and the movers per population. Interestingly, the ratio of inflows over outflows—a crude measure of the attractiveness of a destination for firms—varies widely across cities. Note an immediate pattern: a number of traditional startup hubs are net losers of startups between birth and 5 years of age. Boston, San Francisco, San Jose, and New York City all have more firms move away than move in. The opposite is true of many Sunbelt cities, Chicago, Seattle, Minneapolis, and Denver.

Column 1 of Table 4 reports the relative utility of movers as estimated from the full matrix of startup location choices using the model in Section 3. The rightmost column gives the same rank if we only look at LLCs. Though the ratio of net moves is highly correlated with average relative utility for movers for large cities, in smaller cities, the additional information provided by whether movers in or movers out come from relatively high utility locations is informative.

Note an immediate pattern. High-utility cities are dominated by the Sunbelt (Dallas, Phoenix, Austin, San Antonio, Jacksonville, San Diego) and the New South (Charlotte, Nashville, Atlanta, Jacksonville, Raleigh, Birmingham, Richmond, and Tampa). San Jose, Boston, San Francisco, and New York are all below median cities among CBSAs with a population over one million. In the complete list of cities (Appendix Table A2), note that university towns are particularly likely to show low utility for movers relative to founders.

The fact that “startup hubs” do not have particularly high utility to founders is not due to an idiosyncrasy in how we define “high-growth startups”, as can be seen by investigating

where these firms are born. Appendix Table A3 shows cities by the number of high-growth corporation births per capita. San Jose, Bridgeport, San Francisco, Boulder, Boston, and Durham-Chapel Hill have the highest number, in line with intuition that these are hubs of startup activity (in Bridgeport’s case, for financial sector activity). The cities with the lowest number of high-growth corporation births per capita are Biloxi, Youngstown, Buffalo, Corpus Christi, El Paso, Tucson and Rochester, again in line with expectations. Among the cities with high utility for movers, some also birth many startups (Austin) while others are attractive despite not creating a particularly large number of startups given their size (Seattle, Minneapolis).

Appendix Table A1 constructs the same utility ranking using only LLC moves. Warm-weather “lifestyle” cities loom particularly large: San Diego, Miami, Phoenix, Austin, Los Angeles, and Tampa are all among the top twelve large cities by LLC mover utility. The business centers of the New South - Atlanta, Dallas, Houston, Charlotte - possess much less utility for LLCs than they do for corporations.

Recall that these utility rankings derive from a nonparametric revealed preference model. The average relative utility of cities to startups considering a move has value even without investigating the correlates of those utilities. For instance, as we have noted, it is immediately clear that “the rich do not get richer” when it comes to traditional startup hubs and university towns. It is also immediate that, especially for corporations, the growing metropolitan areas of the New South exert a particular pull.

Before turning to potential explanations of our utility rankings, and the meaning of those explanations for popular theories of startup geography, we should note that firm moves are, empirically, not simply idiosyncratic. Of course, some startup moves are heavily influenced by the idiosyncratic preferences of founders.²⁰ The relative weight on city common utility versus firm idiosyncratic factors, and hence the extent to which city fixed characteristics drives startup location choice, can be directly investigated by considering the pair-wise migration

²⁰For example, Microsoft’s relocation to Seattle appears to be partially influenced by the fact that Bill Gates and Paul Allen wanted to be close to their families.

rates amongst two cities. Taking the model seriously, if the idiosyncratic factor has zero variance, then all firms who move will go to the same city. In contrast, if only idiosyncratic factors matter, then bilateral flows will be identical in each direction. That is, the hypothesis that moves are idiosyncratic directly implies that

$$H_0 : E[\frac{Moves\ In_i}{Moves\ Out_i}] = 1 \quad (1)$$

Table 3 shows ample variation in estimated log-utility of each city, suggesting that bilateral moves are not identical. We implement an empirical version of (1) in Appendix Figure A1, where we report 5000 bootstrap realizations of the mean of Moves In over Moves Out for each CBSA in our sample. A Fisher exact test rejects the hypothesis that moves are idiosyncratic with $p = 0.998$.

Finally, we consider in Figure 8 the relationship between these utilities and an estimate of the vertical quality at birth of firms moving to or from each city. We do this to understand whether differences in the startup quality of firms that arrive versus leave induces sorting across cities. For example, places with high startup costs such as the Bay Area may attract few startups, but they may tend to shed low quality startups while gaining very high quality ones. Note that our primary sample already restricts only to Delaware-registered corporations at birth, so this robustness check is attempting to handle quality differences within a sample that is already highly selected on quality at birth.

To do so, we replicate, within our data, the entrepreneurial quality measure of Guzman and Stern (2020), which maps the founding characteristics of startups before moving to estimated probabilities of reaching an equity outcome such as an IPO or acquisition. Specifically, for all non-movers born before 2012, we run a logit model with a binary measure of equity events as the dependent variable, and observables for whether a firm, close to founding and in its birth location, is a corporation, has a short name, is eponymous, has a patent, has a trademark, has both a patent and a trademark, or is estimated to be part of certain indus-

tries based on firm name.²¹ Estimated quality is then the predicted out of sample probability of this model. We average this value for all movers in and out of a city, and for firms born in a city that do not move.

Panels A and B of Figure 8 plot this relationship for each city with at least five movers. There is only a weak relationship between our estimated city utilities and the quality of movers in or out. If anything, cities with higher estimated utility attract better firms and lose fewer low-quality ones than cities with lower estimated utility, meaning that quality adjustments are unlikely to change our primary estimates. Furthermore, as reported in Appendix Table A4, an OLS regression of the estimated quality of startups that arrive versus depart cities, controlling for the average quality of all firms born in that city, reports a coefficient indistinguishable from zero.²² That is, startup hubs like Silicon Valley do in fact attract high quality startups, but the startups they lose are also disproportionately high quality, and once the average quality of all nonmovers in a city is controlled for, there is essentially no correlation between movers in and out. In short, we do not find evidence that our primary results are driven by quality-based vertical sorting.

5.1 Agglomeration Theory and Startup Mover Utility

Let us now examine what factors drive startup migration. We begin by verbally describing seven well-known theories of startup behavior, then linking those theories to comparative statics between covariates and startup outcomes. We are not claiming to be testing these theories *in toto*, nor to be establishing causality between certain factors and startup migration. Rather, we only wish to investigate in a theoretically-consistent way whether these theories are *first-order* for explaining entrepreneurial migration. Table 5 summarizes the predictions of each model.

First, consider the *Marshallian theory* of agglomeration. This theory emphasizes that firms in the same industry share resources, and these shared resources reduce the cost of

²¹Predictions from this model report an out of sample ROC score of 0.80.

²²The coefficient is 0.01 with a robust standard error of 0.054.

transporting goods, exchanging ideas, and hiring trained workers (Marshall and Marshall, 1920). There is well-identified evidence, at least within manufacturing, that all three factors help explain the coagglomeration patterns of firms (Ellison et al., 2010). If Marshallian concerns drove startup behavior, we should both observe more startups per capita in industrially-specialized cities, and further observe startups moving to even more specialized regions than where they are born. In the context of the model, Marshallian agglomeration implies that the relative utility to movers $\bar{V}$ should be increasing in the industrial specialization of a city, as measured by the Herfindahl-Hirschman index of industrial similarity.

While Marshall focuses on specialization as a source of agglomeration benefits, other writers have argued the opposite, particularly for early-stage and creative business activity. The *Jacobs theory* argues that the underlying idea diversity of a city is what matters to successful business activity (Jacobs, 1969). A modification of this idea is the *nursery city explanation* of (Duranton and Puga, 2001). In the nursery cities model, diverse cities are a good area for innovation but not for scaling companies, leading to net out migration of good companies from college towns and other highly innovative areas. Both models suggest that “idea diverse” cities, whether measured in employment HHI, the share of the population with a university degree, the number of patents per capita, or the share of the population who is a current student, should have a high number of startups per capita. The models differ in whether these cities will be attractive to already-existing startups: the Jacobs explanation should see high mover utility for industrially diverse, high-patent, high-education cities, while the nursery city explanation will see startups move away from these places to specialized cities.

The fourth theory we consider is that of *startup ecosystems*, the focus of Chinitz (1961). This theory argues that “startup ecosystems” provide spillovers especially to small, fast-growing firms from a concentration of labor, capital, specialized services, and culture that is shared across startups. For example, a culture of entrepreneurship causes a positive externality on other startups by deepening the pool of experienced operators. That is, there

are Marshallian agglomerations, at the margin of shared firm age rather than industry. The main prediction of this theory is that high entrepreneurship cities—cities that birth many startups per capita, for whatever region—also should see even higher utility to movers who can take advantage of spillovers from the founders that begin in that city. That is $\bar{V}$ is increasing in startup births per capita.

Next, we move beyond local industry composition and the drivers of firm productivity, to instead consider city amenities that enter the founder’s own utility function. The *amenities explanation* focuses on the fact that while founder utility incorporates startup productivity, it also takes into account the amenity value of the location where the founder lives (Roback, 1982; Albouy, 2016; Guzman, 2018). This interplay between founder utility and firm performance is particularly stark in young firms, where a founder serves as both its principal and its agent. Founders therefore may have high utility from cities with a high urban amenity: sunny days, no rain, comfortable temperatures, and man-made amenities can all factor. Indeed, even growth-oriented startups may weigh these amenities depending on how they enter the utility function. For example, if labor is easier to attract to high-amenity cities, these factors may not be important solely to the founder’s own nonpecuniary interests. A particular amenities-based explanation with great policy purchase draws on the work of Florida (2002). Florida builds on Jacobs and Roback to more concretely emphasize a specific component of the local culture—whether it is bohemian, meaning home to artists and other “creative city” professions—in creating creativity, innovation, and eventual economic growth. The *Bohemian explanation* is therefore that cities with a bohemian culture both create startups and attract startups. Intuitively, amenities-based explanations should be more relevant to the location choice of firms where the founders’ non-pecuniary personal utility is particularly important - in our case, LLCs, which are less growth-oriented, should therefore correlate more strongly to these quality-of-life measures.

Finally, there is a more prosaic *tax policy explanation*. A long line of both research and policy has argued that there is meaningful sensitivity to taxes in the migration choices of

inventors (Moretti and Wilson, 2017; Akcigit et al., 2016) and firms (Fajgelbaum et al., 2019; Slattery, 2018). In our setting, profit maximizing founders would trade off the locational benefits of a region to the amount of tax incentives. As corporations are more growth-oriented, we expect corporations will react more strongly to tax policy on top incomes, while LLCs should react more to taxes lower down the income distribution.

To what extent do we see first-order support for these theories in explaining entrepreneurial migration? We first examine in Table 6 the factors which affect where startups are born. In line with Jacobs and nursery city explanations, more industrially diverse cities birth more high-growth startups per capita, not more industrially specialized cities. The effect of industrial diversity on births per capita is similar for corporations and for LLCs. As suggested by Jacobs, nursery city, startup ecosystem, and Bohemian explanations, more startups are born in highly-educated, Bohemian cities, and in ones with many patents and lots of venture capital investment per capita. All of these factors are of course correlated with income. Broadly, however, the common intuition is correct that startup hubs are creative cities, with many educated workers producing novel inventions.

Where this intuition goes wrong is in assuming that the same pattern holds for startups who move. Table 7 considers whether specialized cities, or diverse, idea-generating cities attract startups. Columns (1) to (4) show that mover utility for corporations is lower in industrially specialized cities and lower in college towns. That is, contra Marshall, startups do not move, in general, to more specialized cities. In particular, college towns, which are good at creating startups, do not attract them, and indeed tend to lose many of the ones they created in the first place. This pattern is particularly evident for corporations rather than LLCs. This may be seen as both support for, and a refutation of, the theory of Jane Jacobs. Startups do indeed prefer diverse, idea-generating cities, but cities which are too specialized in academic work are not attractive.

Table 8 looks at ecosystem effects. There is a weak correlation between cities which birth many startups per capita and those which attract them. Note, however, that the most

well-known startup hubs of San Jose, San Francisco and Boston do not have particularly high mover utility. The amount of VC invested per capita is uncorrelated with mover utility - it is not the case that startups are simply moving to traditional VC hubs. LLCs do see a much stronger link between previous corporate births per capita and their mover utility, perhaps reflecting a combination of the fact that cities with high births per capita are also high-amenity, well-educated cities.

Table 9 turns to amenities. Columns (1) through (4) regress various amenity factors on corporation mover utility. Column (1) incorporates the individual measures of amenity from Albouy (2016), number of degree days, sunshine percentage, distance from water, and latitude. Column (2) follows the standard approach of the urban literature using average home value as a measure of the benefit of living in a location. Column (3) incorporates instead the Albouy (2016) quality of life index, which is a summary measure of locational amenities. Column (4) looks at Florida’s Bohemia index. The point estimate of each type of amenity on mover utility is *negative*, not positive, though none are statistically different from zero. Together, this suggests that high quality startups are not sensitive to the amenities offered by a destination city in their migration choices.

On the other hand, Columns (5) to (8) investigate the same relationships using LLCs. In contrast to the previous results, we find large and significant estimates of city amenities in driving migration choices. Column (5) shows that an important correlate to location choices is the amount of sunshine in a location, a pattern that has been previously observed by other urban research in studying the migration flows of people within the U.S. (e.g. Glaeser, 2011) Column (7) shows that high quality-of-life as measured in Albouy’s composite index is particularly attractive to moving LLCs, while Column (8) shows that LLCs move to Bohemian cities. Comparing the LLC and corporation results suggests an explanation in line with our earlier discussion: the personal utility of founders matters for migrating startups, but especially so for lower-quality startups where the financial upside of relocating

to a less-enjoyable city may not be worth it.²³

We study the relationship of tax incentives to estimated city utility in Table 10. To do so, we consider the relationship of marginal tax rates at multiple levels—the 50th percentile of income, the 95th percentile of income, and the 99th percentile of income—to both corporation and LLC estimates of city utility. We observe two distinct patterns. First, the response to marginal tax rates is substantial across all measures. For corporations, the point estimate of top tax rates on mover utility is negative and strongly significant. For LLCs, median tax rates are most strongly negatively correlated with mover utility. The magnitudes are not small. A ten percentage point lower 99th percentile tax rate lowers mover utility .82 points, roughly half the distance between the highest and lowest utility major cities in Table 4.²⁴

Summing up, industrially diverse and highly-educated cities do indeed produce many startups per capita. However, these startup hubs are not the ones, in general, that attract movers: the rich do not necessarily get richer. Among lower-quality startups, city amenities like sunshine appear important to movers, but these amenities do not attract higher-quality startups, nor do “Bohemian” amenities appear first-order for attracting them. On the other hand, there appears to be an important role of personal income taxes in driving the destination choice of migrants, particularly top income tax rates when it comes to high-quality startups. These results paint a nuanced and interesting picture of the patterns of startup migration across regions, and what they imply about the underlying city attributes that startups value when moving. They further highlight the importance of considering the heterogeneity of firm quality when considering location choices and policy, as we observe

²³For each of the covariates mentioned here, a joint regression was performed including the covariate interacted with an LLC dummy. This tests whether the LLC coefficients are different from those of corporations, not whether they are different from zero. For each of the covariates mentioned above - sunshine, quality of life, and Bohemia, the interaction is significant at the five percent level.

²⁴Transforming the results in the previous two paragraphs into standardized changes, and looking only at point estimates, does not change the qualitative nature of these results. A one standard deviation increase in sunshine percentage causes a roughly .05 standard deviation increase in log-utility for corporations, and .17 for LLCs. A one SD increase in the Albouy quality of life index decreases corporation log-utility by .08, and increases LLC log-utility by .16. A one SD decrease in personal income taxes at the 99th percentile increases corporation log-utility by .2, and LLC log-utility by .085. That is, for corporations, the estimated quantitative importance of variation in tax rates for migration dominates that of even the most important amenities.

significant differences even within the narrow band of high-quality Delaware firms.

We close with an important but significant caveat of our results by emphasizing that they only reflect the utility in the observed choices of startup migrants, and not the value for other types of related populations, such as migrant innovators, migrant small firms, or the local population of startups that does not move. Further, we do not mean to suggest that any of these agglomerative factors are unimportant in contexts other than startup mobility - in many other cases, the importance of these factors is well-established in the literature.

6 Conclusion

Some cities are born lucky: they produce many high-growth startups as a result of their youthful demographics, their technical universities, or spinouts from their keystone firms. After birth, however, 6.6% of those high-growth startups will move before they are five years old. Historically, some of the most important startups moved when young: Facebook, Microsoft, Slack, and Tableau are all prominent examples.

We show that the places which create a lot of startups and the ones that are attractive to movers are not the same. We are able to track startups moves across 36 states making up over 82% of the US population using business registration data. This dataset allows us to capture startups before they ever appear in censuses or other public records, and in a way that is neutral to their industry. Although the usual suspects of San Jose, San Francisco, and Boston produce a wildly disproportionate number of high-growth startups per capita, startup moves are not a case of “the rich get richer”. Comparing the utility of startups when moving to the utility distribution that justifies their original location, it is largely low-tax, not-terribly-Bohemian Sunbelt and southern cities which are more attractive to the highest-quality movers. High-tax “startup hubs” and university towns have more startups move away than move in.

Our method for estimating startup mover utility is wholly nonparametric and based

on revealed preference, using a technique from linear algebra previously applied by Google to identifying important websites, and to compensating differentials in labor economics by Sorkin (2018). This technique allows us to compare cities even when they have very few, or even zero, bilateral moves between them, and even when the econometrician has no a priori knowledge of the covariates which startups consider when planning a move. Using these utility estimates, we find that the most growth-oriented firms, Delaware corporations at birth, appear particularly responsive to tax changes, with idea-heavy and highly-educated cities showing no particular attraction. Firms where the founder’s personal non-pecuniary utility might be thought to be more important in firm location decisions, Delaware LLCs at birth, are highly responsive to quality of life factors like weather when they relocate, while also caring heavily about taxation.

While these estimates are not causal, the results shown here suggest an important focus for spatial entrepreneurship in understanding “startup hubs” as two distinct types of cities: those that create a lot of firms given their population, and those that attract these firms if they choose to leave. It also suggests that college towns and other highly-educated places may not be as advantaged as many think. Although they create many startups, those homegrown firms do not create spillovers sufficient to attract more firms from outside. Indeed, quite the opposite. Many university spinouts leave for the types of cities attractive to businesses of all vintages. Studies of spatial entrepreneurship therefore need to carefully separate factors which birth firms and those which affect the post-migration final locations of those startups.

References

- Jason Abaluck and Abi Adams. Discrete choice models with consideration sets: Identification from asymmetric cross-derivatives, 2016.
- Ufuk Akcigit, Salome Baslandze, and Stefanie Stantcheva. Taxation and the international mobility of inventors. *American Economic Review*, 106(10):2930–81, 2016.
- David Albouy. What are cities worth? land rents, local productivity, and the total value of amenities. *Review of Economics and Statistics*, 98(3):477–487, 2016.
- RJ Andrews, Catherine Fazio, Jorge Guzman, Yupeng Liu, and Scott Stern. The startup cartography projec. *Working Paper*, 2020.
- Catherine Beaudry and Andrea Schifffauerova. Who’s right, marshall or jacobs? the localization versus urbanization debate. *Research policy*, 38(2):318–337, 2009.
- Guido Buenstorf and Steven Klepper. Heritage and agglomeration: the akron tyre cluster revisited. *The Economic Journal*, 119(537):705–733, 2009.
- Christian Catalini, Jorge Guzman, and Scott Stern. Passive versus active growth: Evidence from founder choices and venture capital investment. Technical report, National Bureau of Economic Research, 2019.
- Aaron Chatterji, Edward Glaeser, and William Kerr. Clusters of entrepreneurship and innovation. *Innovation policy and the economy*, 14(1):129–166, 2014.
- Benjamin Chinitz. Contrasts in agglomeration: New york and pittsburgh. *The American Economic Review*, pages 279–289, 1961.
- Annamaria Conti and Jorge Guzman. What is the us comparative advantage in entrepreneurship? evidence from israeli migration to the united states. *Working Paper*, 2019.

- Michael S Dahl and Olav Sorenson. The embedded entrepreneur. *European management review*, 6(3):172–181, 2009.
- Michael S Dahl and Olav Sorenson. Home sweet home: Entrepreneurs’ location choices and the performance of their ventures. *Management science*, 58(6):1059–1071, 2012.
- Ryan Decker, John Haltiwanger, Ron Jarmin, and Javier Miranda. The role of entrepreneurship in us job creation and economic dynamism. *Journal of Economic Perspectives*, 28(3):3–24, 2014.
- Mercedes Delgado, Michael E Porter, and Scott Stern. Clusters and entrepreneurship. *Journal of economic geography*, 10(4):495–518, 2010.
- Gilles Duranton and Diego Puga. Nursery cities: Urban diversity, process innovation, and the life cycle of products. *American Economic Review*, 91(5):1454–1477, 2001.
- Glenn Ellison, Edward L Glaeser, and William R Kerr. What causes industry agglomeration? evidence from coagglomeration patterns. *American Economic Review*, 100(3):1195–1213, 2010.
- Pablo D Fajgelbaum, Eduardo Morales, Juan Carlos Suárez Serrato, and Owen Zidar. State taxes and spatial misallocation. *The Review of Economic Studies*, 86(1):333–376, 2019.
- Richard Florida. Bohemia and economic geography. *Journal of economic geography*, 2(1):55–71, 2002.
- William H Frey. U.s. migration still at historically low levels, census shows, 2017.
- Edward Glaeser. *Triumph of the City*. Pan, 2011.
- Edward L Glaeser and William R Kerr. Local industrial conditions and entrepreneurship: how much of the spatial distribution can we explain? *Journal of Economics & Management Strategy*, 18(3):623–663, 2009.

- Edward L Glaeser, Sari Pekkala Kerr, and William R Kerr. Entrepreneurship and urban growth: An empirical assessment with historical mines. *Review of Economics and Statistics*, 97(2):498–520, 2015.
- Paulo Guimaraes, Octávio Figueirido, and Douglas Woodward. A tractable approach to the firm location decision problem. *Review of Economics and Statistics*, 85(1):201–204, 2003.
- Jorge Guzman. Go west young firm: The value of entrepreneurial migration for startups and founders. *Working Paper*, 2018.
- Jorge Guzman and Scott Stern. Where is silicon valley? *Science*, 347(6222):606–609, 2015.
- Jorge Guzman and Scott Stern. The state of american entrepreneurship: New estimates of the quantity and quality of entrepreneurship for 32 us states, 1988–2014. *American Economic Journal: Economic Policy*, 12(4):212–43, November 2020. doi: 10.1257/pol.20170498. URL <https://www.aeaweb.org/articles?id=10.1257/pol.20170498>.
- John Haltiwanger, Ron S Jarmin, and Javier Miranda. Who creates jobs? small versus large versus young. *Review of Economics and Statistics*, 95(2):347–361, 2013.
- Chang-Tai Hsieh and Enrico Moretti. Housing constraints and spatial misallocation. *American Economic Journal: Macroeconomics*, 11(2):1–39, April 2019.
- Erik Hurst and Benjamin Wild Pugsley. What do small businesses do? *Brookings Papers on Economic Activity*, page 73, 2011.
- Jane Jacobs. Strategies for helping cities. *The American Economic Review*, 59(4):652–656, 1969.
- William R. Kerr and Frederic Robert-Nicoud. Tech clusters. *Journal of Economic Perspectives*, 34(3):50–76, August 2020.
- Alfred Marshall and Mary Paley Marshall. *The economics of industry*. Macmillan and Company, 1920.

- Claudio Michelacci and Olmo Silva. Why so many local entrepreneurs? *The Review of Economics and Statistics*, 89(4):615–633, 2007.
- Enrico Moretti and Daniel J Wilson. The effect of state taxes on the geographical location of top earners: evidence from star scientists. *American Economic Review*, 107(7):1858–1903, 2017.
- Mitsuru Ota and Masahisa Fujita. Communication technologies and spatial organization of multi-unit firms in metropolitan areas. *Regional science and urban economics*, 23(6):695–729, 1993.
- Michael E. Porter. The competitive advantage of nations. *Harvard Business Review*, 68(2):73–93, 1990.
- Jennifer Roback. Wages, rents, and the quality of life. *Journal of political Economy*, 90(6):1257–1278, 1982.
- AnnaLee Saxenian. *Regional networks: industrial adaptation in Silicon Valley and route 128*. Citeseer, 1994.
- Kurt Schmidheiny and Marius Brulhart. On the equivalence of location choice models: Conditional logit, nested logit and poisson. *Journal of Urban Economics*, 69(2), 2011.
- Antoinette Schoar. The divide between subsistence and transformational entrepreneurship. *Innovation policy and the economy*, 10(1):57–81, 2010.
- Scott Shane. Why encouraging more people to become entrepreneurs is bad public policy. *Small business economics*, 33(2):141–149, 2009.
- Cailin Slattery. Bidding for firms: Subsidy competition in the us. *Working Paper*, 2018.
- Isaac Sorkin. Ranking firms using revealed preference. *The quarterly journal of economics*, 133(3):1331–1393, 2018.

Figures & Tables for Entrepreneurial Migration

Figure 1: Estimated Migrant Value of Cities vs Local Ecosystem Strength

A. . Migration Rate vs Local Startups

B.. Migration Rate vs Local Startups per Capita

Figure 2: New Startups Per Capita by MSA

Circle size represents the population of each metropolitan area. Darker circles signify more high-quality startups per capita.

Figure 3: Ratio of Startups Moving In Over Moving Out by MSA

This map plots all cities within the contiguous 48 states in our data. The size of the circle is the number of total movers that move into a metropolitan area. The color of the circle represents the ratio of the moves in over the moves out. Red cities have a higher number of moves in than moves out, while blue cities are the opposite. White cities are about the same.

Figure 4: Net Migration Rates and Taxes

Notes: This figure compares the net migration rate of firms, estimated as the log of the ratio of in moves over out moves, to the average personal income tax rate in that state, estimated by Moretti and Wilson (2017). We observe a large negative correlation between both variables.

Figure 5: 36 Jurisdictions Included in Dataset (82% of US Population)

Figure 6: Migration Rate by Age

Notes: This figure reports the average unconditional probability of moving by age for startups. Most startups move early, but many also do not survive to be considered in the later periods. Our data does not allow us to report the probability of moving conditional on survival.

Figure 7: Migration Rate by Founding Year, 1988-2014

Notes: This figure reports the share of startups by founding year cohort that move within 5 years in our data across states to different CBSAs. Across different subsets of the data, we observe as consistent reduction in the net migration rate of firms in our data, which appear homologous to the observed decline in the cross-state migration of U.S. population.

Figure 8: Estimated Quality of Movers vs Migrant City Utility

A. Quality of Out Migrants and Estimated City Utility

B. Quality of In Migrants and Estimated City Utility

Notes: This figure plots the estimated utility through our method for cities with at least 5 migrants, and the entrepreneurial quality of these migrants following the approach of Guzman & Stern (2020) within our data. Quality is the predicted out of sample probability that a firm achieves a growth event based on founding characteristics before migration. Panel A compares city utility to the average quality of firms that leave a city. The circle size represents the number of movers. Panel B compares city utility to the average quality of firms that arrive into a city. The circle size represents the number of movers. All cities with less than five movers are excluded from the graph. The fitted lines are weighted by the number of movers in or out of the city.

Table 1: Summary Statistics

<i>Panel A: Corporations (N=181663)</i>		
Statistic	Mean	St. Dev.
Incorporation Year	2,001.819	7.470
IPO	0.005	0.071
Acquired	0.026	0.159
Short Name	0.598	0.490
Patent	0.058	0.233
Trademark	0.024	0.153
Moves in 2 years	0.033	0.180
Moves in 5 years	0.056	0.231
<i>Panel B: LLCs (N=237307)</i>		
Statistic	Mean	St. Dev.
Incorporation Year	2,006.634	5.320
IPO	0.0001	0.010
Acquired	0.004	0.065
Short Name	0.377	0.485
Patent	0.015	0.120
Trademark	0.010	0.102
Moves in 2 years	0.018	0.134
Moves in 5 years	0.028	0.166
<i>Panel C: Estimated 5-year U.S. Migration Rates</i>		
Corporations	0.066	
LLCs	0.032	

Table 2: Summary Statistics for Migrant Flows Data

<i>Panel A: State to State Migration Flows</i>			
Statistic	Mean	St. Dev.	N
Number of Corporation Movers	3.701	10.305	1,260
Number of Corporation Movers Cond. on ≥ 1	6.960	13.309	670
Number of LLC Movers	4.155	11.337	1,260
Number of LLC Movers Cond. on ≥ 1	9.519	15.607	550
<i>Panel B: CBSA to CBSA Migration Flows</i>			
Statistic	Mean	St. Dev.	N
Number of Corporation Movers	0.228	1.783	34,040
Number of Corporation Movers Cond. on ≥ 1	3.826	6.927	1,749
Number of LLC Movers	0.153	1.474	34,040
Number of LLC Movers Cond. on ≥ 1	3.689	6.971	1,196

Table 3: Summary Statistics for Metropolitan Areas

Statistic	Mean	St. Dev.	N	Min	Max
Population	978,560	1,935,881	185	73,696	18,641,695
Industry Measures					
Industry Concentration (HHI)	250.3	56.8	185	174.725	518.0
Education and Idea Generation					
Share College Educated	0.247	0.063	185	0.132	0.462
Share University Students	0.058	0.032	185	0.003	0.210
College Town	0.205	0.381	185	0	1
Patents per Thousand Pop	5.982	9.000	185	0.502	84.757
Income Tax					
Income Tax at 50th Percentile	0.107	0.016	185	0.085	0.138
Income Tax at 95th Percentile	0.236	0.023	185	0.199	0.272
Income Tax at 99th Percentile	0.314	0.024	185	0.275	0.347
Venture Capital					
Total VC Invested (mlns)	2,190.80	9,674.264	185	0.000	99,600.0
Number of VC Startups Financed	249	990.9	185	0	9,067
Amenities					
Average Home Value	168,082	78,281	185	82,397	472,353
Cooling Degree Days	1,601.368	1,064.410	185	2	4,628
Heating Degree Days	3,769.665	2,070.925	185	0	10,193
Sunshine Percentage	60.619	9.254	185	0.000	84.059
Inverse Dist. from Water	-1.331	0.937	185	-2.690	0.000
Latitude	36.76	5.620	185	20.875	62.225
Quality of Life	0.007	0.054	185	-0.104	0.208
Creative Industries					
Bohemia	0.391	0.155	185	0.133	1.033
Startup Cartography Project					
Delaware Corporations Registered	952.6	2,699.2	185	3	20,842
Delaware LLCs Registered	1,238.4	4,181.4	185	2	43,382

Table 4: Estimated Utility for Large US Cities (Population over 1 million in 2010)

Log Utility	Rank	CBSA	CBSA Name	Moves In	Moves Out	2010 Pop.	LLC Rank
-2.2455	1	19100	Dallas-Fort Worth-Arlington, TX	453	215	6,426,214	16
-2.3842	2	38060	Phoenix-Mesa-Chandler, AZ	94	53	4,192,887	9
-2.4	3	12420	Austin-Round Rock-Georgetown, TX	166	88	1,716,289	8
-2.436	4	16740	Charlotte-Concord-Gastonia, NC-SC	108	63	2,217,012	29
-2.4716	5	26420	Houston-The Woodlands-Sugar Land, TX	376	205	5,920,416	17
-2.5033	6	42660	Seattle-Tacoma-Bellevue, WA	145	86	3,439,809	32
-2.5524	7	16980	Chicago-Naperville-Elgin, IL-IN-WI	471	311	9,461,105	11
-2.5824	8	41700	San Antonio-New Braunfels, TX	45	26	2,142,508	25
-2.5872	9	33460	Minneapolis-St. Paul-Bloomington, MN-WI	90	59	3,348,859	18
-2.6072	10	27260	Jacksonville, FL	48	35	1,345,596	1
-2.6117	11	34980	Nashville-Davidson-Murfreesboro-Franklin, TN	96	72	1,670,890	10
-2.6395	12	25540	Hartford-East Hartford-Middletown, CT	89	63	1,212,381	6
-2.708	13	12060	Atlanta-Sandy Springs-Alpharetta, GA	363	286	5,286,728	26
-2.7352	14	40060	Richmond, VA	26	19	1,208,101	23
-2.739	15	39580	Raleigh-Cary, NC	98	77	1,130,490	14
-2.7569	16	19740	Denver-Aurora-Lakewood, CO	237	191	2,543,482	30
-2.77	17	13820	Birmingham-Hoover, AL	51	42	1,128,047	33
-2.7925	18	45300	Tampa-St. Petersburg-Clearwater, FL	95	78	2,783,243	4
-2.7928	19	41740	San Diego-Chula Vista-Carlsbad, CA	153	134	3,095,313	2
-2.8196	20	47260	Virginia Beach-Norfolk-Newport News, VA-NC	19	16	1,676,822	35
-2.8666	21	32820	Memphis, TN-MS-AR	35	35	1,324,829	22
-2.8958	22	40900	Sacramento-Roseville-Folsom, CA	32	28	2,149,127	3
-2.898	23	29820	Las Vegas-Henderson-Paradise, NV	62	56	1,951,269	39
-2.9264	24	33100	Miami-Fort Lauderdale-Pompano Beach, FL	349	314	5,564,635	12
-2.9466	25	36740	Orlando-Kissimmee-Sanford, FL	75	68	2,134,411	20
-2.9866	26	17140	Cincinnati, OH-KY-IN	52	55	2,114,580	36
-2.9909	27	18140	Columbus, OH	52	49	1,901,974	38
-3.0023	28	26900	Indianapolis-Carmel-Anderson, IN	47	49	1,887,877	27
-3.0143	29	31080	Los Angeles-Long Beach-Anaheim, CA	507	544	12,828,837	7
-3.0493	30	41940	San Jose-Sunnyvale-Santa Clara, CA	213	238	1,836,911	15
-3.0505	31	28140	Kansas City, MO-KS	74	78	2,009,342	43
-3.0929	32	17460	Cleveland-Elyria, OH	45	52	2,077,240	31
-3.1366	33	14460	Boston-Cambridge-Newton, MA-NH	487	548	4,552,402	24
-3.1437	34	41180	St. Louis, MO-IL	80	90	2,787,701	19
-3.2156	35	41860	San Francisco-Oakland-Berkeley, CA	336	433	4,335,391	13
-3.2365	36	31140	Louisville/Jefferson County, KY-IN	43	58	1,235,708	41
-3.2478	37	38900	Portland-Vancouver-Hillsboro, OR-WA	100	134	2,226,009	37
-3.2753	38	39300	Providence-Warwick, RI-MA	12	14	1,600,852	21
-3.2799	39	40140	Riverside-San Bernardino-Ontario, CA	27	38	4,224,851	5
-3.3593	40	47900	Washington-Arlington-Alexandria, DC-VA-MD-WV	257	362	5,636,232	28
-3.4623	41	35620	New York-Newark-Jersey City, NY-NJ-PA	615	1038	19,567,410	42
-3.4623	42	41620	Salt Lake City, UT	52	72	1,087,873	40
-3.5586	43	35380	New Orleans-Metairie, LA	35	63	1,189,866	45
-3.6558	44	15380	Buffalo-Cheektowaga, NY	11	19	1,135,509	44
-3.8411	45	40380	Rochester, NY	5	11	1,079,671	34

Table 5: Predicted Comparative Statics on Mover Utility and Firm Births

	<i>Drivers of Firm Productivity</i> (Stronger for Corporations)					<i>Drivers of Entrepreneur Utility</i> (Stronger for LLCs)		
	Estimated City Utility	Estimated City Utility	Estimated City Utility	Births	Births	Estimated City Utility	Estimated City Utility	Estimated City Utility
	$\frac{\partial \bar{V}}{\partial (IndSpec)}$	$\frac{\partial \bar{V}}{\partial (IdeaGen)}$	$\frac{\partial \bar{V}}{\partial B}$	$\frac{\partial B}{\partial (IndSpec)}$	$\frac{\partial B}{\partial (IdeaGen)}$	$\frac{\partial \bar{V}}{\partial (QoL)}$	$\frac{\partial \bar{V}}{\partial (Bohemia)}$	$\frac{\partial \bar{V}}{\partial (Taxes)}$
<i>Agglomeration Theories</i>								
Marshall	+			+				
Jacobs	-	+		-	+			
Nursery Cities	+	-		-	+			
Startup Ecosystem			+		+			
Amenities						+		
Bohemia					+		+	
Personal Income Taxes								-

IndSpec refers to measures of Industrial Specialization, such as the Industry-level HHI. IdeaGen refers to measures of idea generation, including the number of patents per capita, the share of population who are university and graduate students, and the share of the population with a university degree. QoL are measures of quality of life. B are the number of new firms born in a region per capita. Taxes are measures of individual income taxes, either at the median or the right tail.

Table 6: What Predicts City Level Firm Births?

	Log(New Firm Births per Capita)									
	Corporations					LLCs				
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Log(CBSA Industry Herfindahl)	-1.037*** (0.327)					-1.395*** (0.422)				
Share College Educated		7.185*** (0.947)					7.283*** (1.141)			
Log(Patents Per Capita)			0.379*** (0.068)					0.315*** (0.077)		
Log(VC Invested +1 / Pop)				0.213*** (0.026)					0.219*** (0.029)	
Bohemia Index					2.772*** (0.376)					3.328*** (0.424)
Observations	185	185	185	185	185	185	185	185	185	185
R ²	0.059	0.279	0.203	0.306	0.253	0.066	0.179	0.088	0.203	0.227

OLS regression with Log(New Firm Births Per Capita) as the dependent variable. Columns (1)-(5) use the number of new corporations per capita, while columns (6)-(10) the number of new LLCs. Robust standard errors in parenthesis. Significance denoted as *p<0.1; **p<0.05; ***p<0.01

Table 7: Jacobs and Marshall: Do Specialized Cities, or Diverse, Idea-Generating Cities Attract Startups?

	City Utility							
	Corporations				LLCs			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
HHI	-0.003*** (0.001)				-0.002 (0.002)			
Share College Educated		-0.510 (0.864)				1.711 (1.108)		
Log(Patents Per Capita)			0.024 (0.060)				0.121** (0.061)	
Share Higher Ed. Students				-4.160* (2.428)				3.863 (2.968)
Observations	185	185	185	185	126	126	126	126
R ²	0.060	0.002	0.001	0.028	0.023	0.018	0.024	0.012

OLS regression with City Utility as the dependent variable. City utility is our estimated measure from the underlying graph of moves across cities in the United States. Columns (1)- (4) use the utility estimated through the moves of corporations registered under Delaware jurisdiction (but domiciled anywhere in the U.S.). Columns (5)-(8) use the utility estimated through the moves of LLCs registered under Delaware jurisdiction. We are able to estimate underlying utility for 183 cities using corporations, and 123 cities using LLCs. CBSA Industry Herfindahl is estimated as the total industry share captured by the top 50 industries (by employment) in the CBSA. Share College Educated is the share of individuals who have at least some college in the CBSA. Share Higher Ed. Students is the share of individuals who are studying in a university, including both undergraduate and graduate students. Robust standard errors in parenthesis. Significance denoted as *p<0.1; **p<0.05; ***p<0.01

Table 8: Ecosystems: Does an Existing Startup Ecosystem Attract More Startups?

	City Utility					
	Corporations			LLCs		
	(1)	(2)	(3)	(4)	(5)	(6)
Log(Corporation Births / Pop.)	0.133*			0.322***		
	(0.070)			(0.066)		
Log(LLC Births / Pop.)		0.065			0.318***	
		(0.051)			(0.064)	
Log(Venture Capital Invested +1 / Pop.)			0.017			0.060*
			(0.026)			(0.030)
Observations	185	185	185	126	126	126
R ²	0.021	0.008	0.002	0.116	0.150	0.026

OLS regression with City Utility as the dependent variable. City utility is our estimated measure from the underlying graph of moves across cities in the United States. Columns (1)- (3) use the utility estimated through the moves of corporations registered under Delaware jurisdiction (but domiciled anywhere in the U.S.). Columns (4)-(6) use the utility estimated through the moves of LLCs registered under Delaware jurisdiction. We are able to estimate underlying utility for 183 cities using corporations, and 123 cities using LLCs. *Venture Capital Invested* is the total dollar amount of VC investment in this city between 1988 and 2014, via Crunchbase. Robust standard errors in parenthesis. Significance denoted as *p<0.1; **p<0.05; ***p<0.01

Table 9: Amenities: Do Local Amenties Correlate to Estimated City Utility?

	City Utility							
	Corporations				LLCs			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Cooling Degree Days	0.00002 (0.0001)				-0.00004 (0.0001)			
Heating Degree Days	0.00004 (0.0001)				-0.0002* (0.0001)			
Sunshine Percentage	0.005 (0.007)				0.017*** (0.006)			
Inverse Dist. from Water	0.013 (0.066)				0.170** (0.076)			
Latitude	-0.030 (0.035)				0.053* (0.031)			
Average Home Value		-0.071 (0.068)				0.227*** (0.088)		
Quality of Life Index			-1.497 (1.068)				3.022** (1.505)	
Bohemia				-0.404 (0.397)				0.891* (0.499)
Observations	185	185	185	185	126	126	126	126
R ²	0.034	0.005	0.010	0.006	0.114	0.058	0.045	0.032

OLS regression with city utility as the dependent variable. City utility is our estimated measure from the underlying graph of moves across cities in the United States. Columns (1)- (4) use the utility estimated through the moves of corporations registered under Delaware jurisdiction (but domiciled anywhere in the U.S.). Columns (5)-(8) use the utility estimated through the moves of LLCs registered under Delaware jurisdiction. We are able to estimate underlying utility for 183 cities using corporations, and 123 cities using LLCs. Cooling and Heating Degree Days are the cumulative gaps in days times temperature throughout a year where cooling or heating is needed as the temperature deviates from 65F. Inv. Distance to Water is the log of 1 over the distance (in miles) to the coast or to a Great Lake. Sunshine Percentage is the share of sunny days in a region. Inverse Dist. to Water is the log of the inverse distance to either the ocean or the Great Lakes. Average Home Value is denoted in units of 100,000 dollars. Quality of Life Index uses Albouy (2007) updated estimates of summmary quality of life taking into account these amenities. Bohemia is estimated is estimated by replicating the methodology in Florida (2001): the location quotient of the incidence of professionals in bohemian occupations. These occupations are: authors; designers; musicians and composers; actors and directors; craft-artists, painters, sculptors, and artist printmakers; photographers; dancers; and artists, performers, and related workers. Robust standard errors in parenthesis. Significance denoted as *p<0.1; **p<0.05; ***p<0.01

Table 10: Taxes: Do State-Level Personal Income Taxes Correlate to Estimated City Utility?

	City Utility					
	Corporations			LLCs		
	(1)	(2)	(3)	(4)	(5)	(6)
Personal Income Tax at 50th Perc	−6.683* (3.548)			−13.271*** (3.630)		
Personal Income Tax at 95th Perc		−8.185*** (2.305)			−4.858* (2.600)	
Personal Income Tax at 99th Perc			−8.223*** (2.190)			−3.541 (2.600)
Observations	185	185	185	126	126	126
R ²	0.018	0.058	0.063	0.077	0.022	0.013

OLS regression with city utility as the dependent variable. City utility is our estimated measure from the underlying graph of moves across cities in the United States. Columns 1-3 use the utility estimated through the moves of corporations registered under Delaware jurisdiction (but domiciled anywhere in the U.S.).

Columns 4-6 use the utility estimated through the moves of LLCs registered under Delaware jurisdiction. Personal income tax estimates are taken from Moretti and Wilson (2017), who estimate state-level taxes for all U.S. at different points of the income distribution. Robust standard errors in parenthesis. Significance denoted as *p<0.1; **p<0.05; ***p<0.01

Appendix

Figure A1: Are Moves Idiosyncratic?

Mean of Moves Out / Moves In.
5000 Bootstrap Realizations

Fisher Exact p-value for Ratio Larger than One: $p=1$

Table A1: Estimated Utility for Large US Cities (Population over 1 million in 2010) by LLC Rank.

Log Utility	Rank	CBSA	CBSA Name	LLC Moves In	LLC Moves Out	2010 Pop.	LLC Rank	Log Utility LLC
-2.607	10	27260	Jacksonville, FL	49	19	1,345,596	1	-1.76
-2.793	19	41740	San Diego-Chula Vista-Carlsbad, CA	217	77	3,095,313	2	-1.817
-2.896	22	40900	Sacramento-Roseville-Folsom, CA	38	15	2,149,127	3	-1.885
-2.793	18	45300	Tampa-St. Petersburg-Clearwater, FL	88	42	2,783,243	4	-1.959
-3.28	39	40140	Riverside-San Bernardino-Ontario, CA	23	10	4,224,851	5	-1.965
-2.639	12	25540	Hartford-East Hartford-Middletown, CT	29	15	1,212,381	6	-2.012
-3.014	29	31080	Los Angeles-Long Beach-Anaheim, CA	663	374	12,828,837	7	-2.227
-2.4	3	12420	Austin-Round Rock-Georgetown, TX	96	74	1,716,289	8	-2.302
-2.384	2	38060	Phoenix-Mesa-Chandler, AZ	47	32	4,192,887	9	-2.309
-2.612	11	34980	Nashville-Davidson-Murfreesboro-Franklin, TN	86	61	1,670,890	10	-2.334
-2.552	7	16980	Chicago-Naperville-Elgin, IL-IN-WI	341	261	9,461,105	11	-2.365
-2.926	24	33100	Miami-Fort Lauderdale-Pompano Beach, FL	339	239	5,564,635	12	-2.366
-3.216	35	41860	San Francisco-Oakland-Berkeley, CA	242	154	4,335,391	13	-2.374
-2.739	15	39580	Raleigh-Cary, NC	33	24	1,130,490	14	-2.383
-3.049	30	41940	San Jose-Sunnyvale-Santa Clara, CA	53	35	1,836,911	15	-2.392
-2.245	1	19100	Dallas-Fort Worth-Arlington, TX	324	241	6,426,214	16	-2.404
-2.472	5	26420	Houston-The Woodlands-Sugar Land, TX	282	184	5,920,416	17	-2.407
-2.587	9	33460	Minneapolis-St. Paul-Bloomington, MN-WI	82	65	3,348,859	18	-2.437
-3.144	34	41180	St. Louis, MO-IL	8	9	2,787,701	19	-2.437
-2.947	25	36740	Orlando-Kissimmee-Sanford, FL	73	58	2,134,411	20	-2.472
-3.275	38	39300	Providence-Warwick, RI-MA	23	21	1,600,852	21	-2.521
-2.867	21	32820	Memphis, TN-MS-AR	43	41	1,324,829	22	-2.554
-2.735	14	40060	Richmond, VA	22	23	1,208,101	23	-2.691
-3.137	33	14460	Boston-Cambridge-Newton, MA-NH	253	253	4,552,402	24	-2.696
-2.582	8	41700	San Antonio-New Braunfels, TX	25	20	2,142,508	25	-2.699
-2.708	13	12060	Atlanta-Sandy Springs-Alpharetta, GA	217	239	5,286,728	26	-2.795
-3.002	28	26900	Indianapolis-Carmel-Anderson, IN	31	47	1,887,877	27	-2.842
-3.359	40	47900	Washington-Arlington-Alexandria, DC-VA-MD-WV	140	181	5,636,232	28	-2.878
-2.436	4	16740	Charlotte-Concord-Gastonia, NC-SC	65	83	2,217,012	29	-2.943
-2.757	16	19740	Denver-Aurora-Lakewood, CO	114	177	2,543,482	30	-2.966
-3.093	32	17460	Cleveland-Elyria, OH	39	50	2,077,240	31	-2.972
-2.503	6	42660	Seattle-Tacoma-Bellevue, WA	43	76	3,439,809	32	-3.13
-2.77	17	13820	Birmingham-Hoover, AL	36	51	1,128,047	33	-3.131
-3.841	45	40380	Rochester, NY	2	4	1,079,671	34	-3.136
-2.82	20	47260	Virginia Beach-Norfolk-Newport News, VA-NC	9	16	1,676,822	35	-3.171
-2.987	26	17140	Cincinnati, OH-KY-IN	24	43	2,114,580	36	-3.281
-3.248	37	38900	Portland-Vancouver-Hillsboro, OR-WA	46	100	2,226,009	37	-3.324
-2.991	27	18140	Columbus, OH	27	42	1,901,974	38	-3.345
-2.898	23	29820	Las Vegas-Henderson-Paradise, NV	33	78	1,951,269	39	-3.41
-3.462	42	41620	Salt Lake City, UT	22	47	1,087,873	40	-3.418
-3.236	36	31140	Louisville/Jefferson County, KY-IN	21	45	1,235,708	41	-3.47
-3.462	41	35620	New York-Newark-Jersey City, NY-NJ-PA	263	731	19,567,410	42	-3.557
-3.051	31	28140	Kansas City, MO-KS	18	36	2,009,342	43	-3.614
-3.656	44	15380	Buffalo-Cheektowaga, NY	1	4	1,135,509	44	-3.755
-3.559	43	35380	New Orleans-Metairie, LA	14	66	1,189,866	45	-3.941

Table A2: Estimated Utility. Full List.

Log Utility	Rank	CBSA	CBSA Name	Moves In	Moves Out	2010 Pop.	LLC Rank	Log Utility LLC
-1.065	1	37900	Peoria, IL	12	2	379,186	NA	
-1.33	2	13980	Blacksburg-Christiansburg, VA	7	1	178,237	NA	
-1.444	3	13140	Beaumont-Port Arthur, TX	5	1	403,190	NA	
-1.546	4	39460	Punta Gorda, FL	4	1	159,978	NA	
-1.605	5	47380	Waco, TX	4	1	252,772	NA	
-1.705	6	15500	Burlington, NC	13	4	151,131	NA	
-1.784	7	21780	Evansville, IN-KY	12	5	311,552	122	-5.063
-1.807	8	40220	Roanoke, VA	8	3	308,707	NA	
-2.045	9	15940	Canton-Massillon, OH	7	3	404,422	NA	
-2.14	10	44420	Staunton, VA	2	2	118,502	NA	
-2.144	11	25860	Hickory-Lenoir-Morganton, NC	7	3	365,497	NA	
-2.153	12	27740	Johnson City, TN	6	2	198,716	NA	
-2.172	13	49180	Winston-Salem, NC	12	6	640,595	108	-3.693
-2.175	14	24540	Greeley, CO	4	2	252,825	NA	
-2.2	15	15260	Brunswick, GA	4	1	112,370	29	-2.33
-2.21	16	34940	Naples-Marco Island, FL	11	5	321,520	25	-2.304
-2.221	17	46060	Tucson, AZ	13	6	980,263	NA	
-2.241	18	20020	Dothan, AL	6	3	145,639	NA	
-2.245	19	19100	Dallas-Fort Worth-Arlington, TX	453	215	6,426,214	39	-2.404
-2.322	20	48620	Wichita, KS	19	10	630,919	23	-2.299
-2.349	21	33740	Monroe, LA	5	3	176,441	96	-3.438
-2.353	22	23540	Gainesville, FL	9	4	264,275	52	-2.614
-2.384	23	38060	Phoenix-Mesa-Chandler, AZ	94	53	4,192,887	27	-2.309
-2.4	24	12420	Austin-Round Rock-Georgetown, TX	166	88	1,716,289	24	-2.302
-2.403	25	42340	Savannah, GA	10	6	347,611	86	-3.297
-2.41	26	37860	Pensacola-Ferry Pass-Brent, FL	5	6	448,991	8	-1.856
-2.411	27	19300	Daphne-Fairhope-Foley, AL	3	2	182,265	NA	
-2.436	28	16740	Charlotte-Concord-Gastonia, NC-SC	108	63	2,217,012	69	-2.943
-2.441	29	19340	Davenport-Moline-Rock Island, IA-IL	1	3	379,690	NA	
-2.46	30	45820	Topeka, KS	8	4	233,870	12	-1.981
-2.472	31	26420	Houston-The Woodlands-Sugar Land, TX	376	205	5,920,416	40	-2.407
-2.476	32	42100	Santa Cruz-Watsonville, CA	9	6	262,382	NA	
-2.483	33	15980	Cape Coral-Fort Myers, FL	21	12	618,754	17	-2.224
-2.503	34	42660	Seattle-Tacoma-Bellevue, WA	145	86	3,439,809	76	-3.13
-2.542	35	33860	Montgomery, AL	12	8	374,536	53	-2.65
-2.552	36	16980	Chicago-Naperville-Elgin, IL-IN-WI	471	311	9,461,105	34	-2.365
-2.568	37	31340	Lynchburg, VA	4	3	252,634	NA	
-2.582	38	41700	San Antonio-New Braunfels, TX	45	26	2,142,508	56	-2.699
-2.587	39	33660	Mobile, AL	7	6	412,992	114	-3.844
-2.587	40	33460	Minneapolis-St. Paul-Bloomington, MN-WI	90	59	3,348,859	41	-2.437
-2.607	41	27260	Jacksonville, FL	48	35	1,345,596	4	-1.76
-2.612	42	34980	Nashville-Davidson-Murfreesboro-Franklin, TN	96	72	1,670,890	30	-2.334
-2.62	43	12700	Barnstable Town, MA	11	9	215,888	123	-5.063
-2.639	44	35980	Norwich-New London, CT	6	4	274,055	NA	
-2.639	45	25540	Hartford-East Hartford-Middletown, CT	89	63	1,212,381	15	-2.012
-2.644	46	18880	Crestview-Fort Walton Beach-Destin, FL	5	3	235,865	26	-2.307
-2.644	47	27620	Jefferson City, MO	4	4	149,807	13	-1.983
-2.648	48	12020	Athens-Clarke County, GA	3	1	192,541	NA	
-2.676	49	36100	Ocala, FL	3	2	331,298	NA	
-2.702	50	41100	St. George, UT	3	2	138,115	NA	
-2.708	51	12060	Atlanta-Sandy Springs-Alpharetta, GA	363	286	5,286,728	62	-2.795
-2.72	52	44700	Stockton, CA	3	3	685,306	NA	
-2.735	53	40060	Richmond, VA	26	19	1,208,101	54	-2.691
-2.739	54	39580	Raleigh-Cary, NC	98	77	1,130,490	37	-2.383
-2.746	55	42200	Santa Maria-Santa Barbara, CA	11	12	423,895	20	-2.265
-2.753	56	10420	Akron, OH	22	17	703,200	61	-2.791
-2.757	57	19740	Denver-Aurora-Lakewood, CO	237	191	2,543,482	70	-2.966
-2.758	58	14860	Bridgeport-Stamford-Norwalk, CT	242	174	916,829	14	-2.01
-2.765	59	40420	Rockford, IL	4	4	349,431	NA	
-2.766	60	10740	Albuquerque, NM	51	42	887,077	110	-3.715
-2.77	61	13820	Birmingham-Hoover, AL	51	42	1,128,047	77	-3.131
-2.778	62	18580	Corpus Christi, TX	3	2	428,185	NA	
-2.793	63	45300	Tampa-St. Petersburg-Clearwater, FL	95	78	2,783,243	10	-1.959
-2.793	64	41740	San Diego-Chula Vista-Carlsbad, CA	153	134	3,095,313	6	-1.817
-2.804	65	22180	Fayetteville, NC	4	3	366,383	NA	
-2.804	66	21340	El Paso, TX	6	4	804,123	31	-2.344
-2.82	67	47260	Virginia Beach-Norfolk-Newport News, VA-NC	19	16	1,676,822	80	-3.171
-2.847	68	22660	Fort Collins, CO	8	6	299,630	63	-2.814
-2.85	69	44060	Spokane-Spokane Valley, WA	4	2	527,753	57	-2.746
-2.866	70	43780	South Bend-Mishawaka, IN-MI	5	4	319,224	58	-2.749
-2.867	71	32820	Memphis, TN-MS-AR	35	35	1,324,829	50	-2.554
-2.869	72	34900	Napa, CA	7	8	136,484	1	-1.549
-2.87	73	24660	Greensboro-High Point, NC	26	22	723,801	73	-3.07
-2.873	74	12540	Bakersfield, CA	4	3	839,631	NA	
-2.875	75	31420	Macon-Bibb County, GA	2	3	232,293	NA	
-2.891	76	21140	Elkhart-Goshen, IN	5	6	197,559	NA	
-2.896	77	40900	Sacramento-Roseville-Folsom, CA	32	28	2,149,127	9	-1.885
-2.898	78	29820	Las Vegas-Henderson-Paradise, NV	62	56	1,951,269	92	-3.41
-2.907	79	35840	North Port-Sarasota-Bradenton, FL	28	20	702,281	16	-2.145
-2.917	80	40580	Rocky Mount, NC	3	4	152,392	NA	
-2.917	81	29740	Las Cruces, NM	4	4	209,233	NA	
-2.921	82	44140	Springfield, MA	11	11	621,570	22	-2.296
-2.926	83	33100	Miami-Fort Lauderdale-Pompano Beach, FL	349	314	5,564,635	35	-2.366
-2.936	84	20500	Durham-Chapel Hill, NC	55	52	504,357	49	-2.55

-2.947	85	36740	Orlando-Kissimmee-Sanford, FL	75	68	2,134,411	43	-2.472
-2.949	86	35300	New Haven-Milford, CT	37	33	862,477	59	-2.776
-2.957	87	45780	Toledo, OH	12	9	610,001	91	-3.377
-2.96	88	14500	Boulder, CO	82	79	294,567	60	-2.785
-2.973	89	45220	Tallahassee, FL	6	6	367,413	19	-2.257
-2.987	90	17140	Cincinnati, OH-KY-IN	52	55	2,114,580	85	-3.281
-2.988	91	37460	Panama City, FL	2	2	184,715	NA	
-2.99	92	16860	Chattanooga, TN-GA	14	14	528,143	95	-3.418
-2.991	93	18140	Columbus, OH	52	49	1,901,974	88	-3.345
-3.002	94	26900	Indianapolis-Carmel-Anderson, IN	47	49	1,887,877	64	-2.842
-3.01	95	38340	Pittsfield, MA	2	3	131,219	7	-1.847
-3.014	96	31080	Los Angeles-Long Beach-Anaheim, CA	507	544	12,828,837	18	-2.227
-3.025	97	31700	Manchester-Nashua, NH	32	31	400,721	109	-3.7
-3.049	98	41940	San Jose-Sunnyvale-Santa Clara, CA	213	238	1,836,911	38	-2.392
-3.051	99	28140	Kansas City, MO-KS	74	78	2,009,342	104	-3.614
-3.074	100	14740	Bremerton-Silverdale-Port Orchard, WA	3	3	251,133	NA	
-3.074	101	28940	Knoxville, TN	17	20	837,571	33	-2.365
-3.084	102	49660	Youngstown-Warren-Boardman, OH-PA	4	4	565,773	105	-3.655
-3.086	103	28700	Kingsport-Bristol, TN-VA	3	4	309,544	NA	
-3.093	104	17460	Cleveland-Elyria, OH	45	52	2,077,240	71	-2.972
-3.12	105	30340	Lewiston-Auburn, ME	3	3	107,702	NA	
-3.125	106	38860	Portland-South Portland, ME	22	25	514,098	74	-3.084
-3.137	107	14460	Boston-Cambridge-Newton, MA-NH	487	548	4,552,402	55	-2.696
-3.141	108	29180	Lafayette, LA	10	18	466,750	107	-3.69
-3.144	109	41180	St. Louis, MO-IL	80	90	2,787,701	42	-2.437
-3.155	110	13780	Binghamton, NY	2	2	251,725	NA	
-3.173	111	44100	Springfield, IL	3	3	210,170	NA	
-3.203	112	29460	Lakeland-Winter Haven, FL	5	5	602,095	66	-2.902
-3.212	113	19430	Dayton-Kettering, OH	9	15	799,232	113	-3.769
-3.212	114	39900	Reno, NV	17	19	425,417	100	-3.504
-3.216	115	41860	San Francisco-Oakland-Berkeley, CA	336	433	4,335,391	36	-2.374
-3.228	116	41500	Salinas, CA	4	8	415,057	NA	
-3.236	117	31140	Louisville/Jefferson County, KY-IN	43	58	1,235,708	98	-3.47
-3.243	118	19660	Deltona-Daytona Beach-Ormond Beach, FL	10	10	590,289	21	-2.272
-3.248	119	38900	Portland-Vancouver-Hillsboro, OR-WA	100	134	2,226,009	87	-3.324
-3.253	120	25060	Gulfport-Biloxi, MS	7	11	370,702	84	-3.232
-3.264	121	49340	Worcester, MA-CT	23	28	916,980	32	-2.359
-3.274	122	16820	Charlottesville, VA	7	7	218,705	103	-3.61
-3.275	123	39300	Providence-Warwick, RI-MA	12	14	1,600,852	47	-2.521
-3.28	124	40140	Riverside-San Bernardino-Ontario, CA	27	38	4,224,851	11	-1.965
-3.294	125	14540	Bowling Green, KY	3	6	158,599	72	-3.007
-3.333	126	26380	Houma-Thibodaux, LA	2	3	208,178	124	-5.063
-3.338	127	37100	Oxnard-Thousand Oaks-Ventura, CA	13	19	823,318	68	-2.935
-3.357	128	11700	Asheville, NC	7	10	424,858	28	-2.313
-3.359	129	47900	Washington-Arlington-Alexandria, DC-VA-MD-WV	257	362	5,636,232	65	-2.878
-3.364	130	31740	Manhattan, KS	2	4	92,719	NA	
-3.37	131	42680	Sebastian-Vero Beach, FL	3	5	138,028	3	-1.622
-3.428	132	46520	Urban Honolulu, HI	16	26	953,207	111	-3.751
-3.431	133	17980	Columbus, GA-AL	2	3	294,865	125	-5.063
-3.462	134	35620	New York-Newark-Jersey City, NY-NJ-PA	615	1038	19,567,410	101	-3.557
-3.462	135	41620	Salt Lake City, UT	52	72	1,087,873	94	-3.418
-3.478	136	13460	Bend, OR	5	8	157,733	106	-3.668
-3.481	137	24420	Grants Pass, OR	2	3	82,713	NA	
-3.512	138	17820	Colorado Springs, CO	16	27	645,613	121	-4.644
-3.526	139	19460	Decatur, AL	2	3	153,829	NA	
-3.559	140	35380	New Orleans-Metairie, LA	35	63	1,189,866	117	-3.941
-3.623	141	21660	Eugene-Springfield, OR	7	10	351,715	81	-3.213
-3.624	142	27140	Jackson, MS	13	28	567,122	102	-3.601
-3.631	143	14020	Bloomington, IN	3	3	159,549	NA	
-3.632	144	26580	Huntington-Ashland, WV-KY-OH	2	4	364,908	NA	
-3.639	145	45060	Syracuse, NY	2	3	662,577	NA	
-3.642	146	26620	Huntsville, AL	9	19	417,593	97	-3.462
-3.652	147	14260	Boise City, ID	18	33	616,561	46	-2.51
-3.656	148	15380	Buffalo-Cheektowaga, NY	11	19	1,135,509	112	-3.755
-3.659	149	42020	San Luis Obispo-Paso Robles, CA	3	5	269,637	2	-1.556
-3.665	150	37340	Palm Bay-Melbourne-Titusville, FL	10	15	543,376	48	-2.524
-3.669	151	44180	Springfield, MO	4	5	436,712	NA	
-3.732	152	39340	Provo-Orem, UT	22	43	526,810	93	-3.416
-3.767	153	32780	Medford, OR	3	4	203,206	NA	
-3.781	154	12260	Augusta-Richmond County, GA-SC	2	4	564,873	NA	
-3.799	155	16580	Champaign-Urbana, IL	3	9	231,891	NA	
-3.841	156	40380	Rochester, NY	5	11	1,079,671	78	-3.136
-3.843	157	33260	Midland, TX	2	3	141,671	126	-5.063
-3.856	158	17660	Coeur d'Alene, ID	2	6	138,494	5	-1.816
-3.902	159	12940	Baton Rouge, LA	9	25	802,484	120	-4.34
-3.905	160	36260	Ogden-Clearfield, UT	5	11	597,159	44	-2.49
-3.938	161	26140	Homosassa Springs, FL	2	2	141,236	NA	
-3.941	162	41420	Salem, OR	2	5	390,738	82	-3.215
-3.947	163	23060	Fort Wayne, IN	2	8	416,257	83	-3.225
-4.037	164	48900	Wilmington, NC	3	7	254,884	99	-3.498
-4.12	165	10580	Albany-Schenectady-Troy, NY	8	29	870,716	115	-3.86
-4.149	166	29940	Lawrence, KS	1	4	110,826	NA	
-4.203	167	12620	Bangor, ME	1	3	153,923	NA	
-4.216	168	42220	Santa Rosa-Petaluma, CA	3	12	483,878	45	-2.501
-4.225	169	30460	Lexington-Fayette, KY	10	34	472,099	90	-3.372
-4.228	170	23580	Gainesville, GA	1	4	179,684	NA	
-4.297	171	38940	Port St. Lucie, FL	3	13	424,107	67	-2.926

-4.359	172	15540	Burlington-South Burlington, VT	12	48	211,261	79	-3.155
-4.487	173	17860	Columbia, MO	1	5	162,642	NA	
-4.522	174	29200	Lafayette-West Lafayette, IN	1	5	201,789	NA	
-4.696	175	25620	Hattiesburg, MS	1	4	142,842	118	-4.083
-4.796	176	11260	Anchorage, AK	2	8	380,821	75	-3.088
-4.837	177	43340	Shreveport-Bossier City, LA	2	8	439,811	119	-4.116
-4.983	178	22140	Farmington, NM	1	4	130,044	NA	
-5.017	179	12100	Atlantic City-Hammonton, NJ	1	8	274,549	51	-2.575
-5.16	180	33700	Modesto, CA	1	6	514,453	NA	
-5.373	181	10540	Albany-Lebanon, OR	0	6	116,672	NA	
-5.373	182	23420	Fresno, CA	0	5	930,450	89	-3.357
-5.373	183	27060	Ithaca, NY	0	7	101,564	NA	
-5.373	184	27980	Kahului-Wailuku-Lahaina, HI	1	6	154,924	116	-3.891
-5.373	185	46220	Tuscaloosa, AL	0	4	230,162	NA	

Table A3: Number of new Delaware Corporations and LLCs by city.

CBSA	2010 Pop.	CBSA Name	Corporations	LLCs	Corps / Pop	LLCs / Pop
41940	1,836,911	San Jose-Sunnyvale-Santa Clara, CA	8,996	4,187	4.8974	2.2794
14860	916,829	Bridgeport-Stamford-Norwalk, CT	3,854	6,428	4.2036	7.0111
41860	4,335,391	San Francisco-Oakland-Berkeley, CA	16,580	18,748	3.8243	4.3244
14500	294,567	Boulder, CO	1,103	685	3.7445	2.3254
14460	4,552,402	Boston-Cambridge-Newton, MA-NH	12,685	15,571	2.7864	3.4204
20500	504,357	Durham-Chapel Hill, NC	1,055	587	2.0918	1.1639
31080	12,828,837	Los Angeles-Long Beach-Anaheim, CA	20,842	43,382	1.6246	3.3816
33100	5,564,635	Miami-Fort Lauderdale-Pompano Beach, FL	8,678	12,503	1.5595	2.2469
41740	3,095,313	San Diego-Chula Vista-Carlsbad, CA	4,793	8,445	1.5485	2.7283
12420	1,716,289	Austin-Round Rock-Georgetown, TX	2,380	3,431	1.3867	1.9991
47900	5,636,232	Washington-Arlington-Alexandria, DC-VA-MD-WV	7,603	5,278	1.349	0.9364
34900	136,484	Napa, CA	177	400	1.2969	2.9307
39580	1,130,490	Raleigh-Cary, NC	1,420	852	1.2561	0.7537
15540	211,261	Burlington-South Burlington, VT	262	143	1.2402	0.6769
42680	138,028	Sebastian-Vero Beach, FL	158	150	1.1447	1.0867
19740	2,543,482	Denver-Aurora-Lakewood, CO	2,711	3,478	1.0659	1.3674
42200	423,895	Santa Maria-Santa Barbara, CA	440	846	1.038	1.9958
12700	215,888	Barnstable Town, MA	220	294	1.019	1.3618
25540	1,212,381	Hartford-East Hartford-Middletown, CT	1,225	845	1.0104	0.697
42100	262,382	Santa Cruz-Watsonville, CA	260	111	0.9909	0.423
31700	400,721	Manchester-Nashua, NH	395	163	0.9857	0.4068
26420	5,920,416	Houston-The Woodlands-Sugar Land, TX	5,598	8,298	0.9455	1.4016
12060	5,286,728	Atlanta-Sandy Springs-Alpharetta, GA	4,928	5,139	0.9321	0.9721
35300	862,477	New Haven-Milford, CT	768	439	0.8905	0.509
19100	6,426,214	Dallas-Fort Worth-Arlington, TX	5,712	10,878	0.8889	1.6928
36740	2,134,411	Orlando-Kissimmee-Sanford, FL	1,884	2,868	0.8827	1.3437
35840	702,281	North Port-Sarasota-Bradenton, FL	610	647	0.8686	0.9213
41620	1,087,873	Salt Lake City, UT	904	947	0.831	0.8705
38340	131,219	Pittsfield, MA	107	193	0.8154	1.4708
16980	9,461,105	Chicago-Naperville-Elgin, IL-IN-WI	7,400	14,257	0.7821	1.5069
39340	526,810	Provo-Orem, UT	402	376	0.7631	0.7137
34940	321,520	Naples-Marco Island, FL	245	406	0.762	1.2628
45300	2,783,243	Tampa-St. Petersburg-Clearwater, FL	2,044	2,603	0.7344	0.9352
35620	19,567,410	New York-Newark-Jersey City, NY-NJ-PA	13,738	14,370	0.7021	0.7344
30460	472,099	Lexington-Fayette, KY	326	300	0.6905	0.6355
49340	916,980	Worcester, MA-CT	614	533	0.6696	0.5813
27260	1,345,596	Jacksonville, FL	861	993	0.6399	0.738
15980	618,754	Cape Coral-Fort Myers, FL	395	508	0.6384	0.821
42220	483,878	Santa Rosa-Petaluma, CA	308	425	0.6365	0.8783
26620	417,593	Huntsville, AL	265	260	0.6346	0.6226
37340	543,376	Palm Bay-Melbourne-Titusville, FL	340	277	0.6257	0.5098
13820	1,128,047	Birmingham-Hoover, AL	698	1,898	0.6188	1.6826
38940	424,107	Port St. Lucie, FL	262	259	0.6178	0.6107
16740	2,217,012	Charlotte-Concord-Gastonia, NC-SC	1,368	2,430	0.617	1.0961
37100	823,318	Oxnard-Thousand Oaks-Ventura, CA	486	461	0.5903	0.5599
34980	1,670,890	Nashville-Davidson-Murfreesboro-Franklin, TN	968	1,575	0.5793	0.9426
41500	415,057	Salinas, CA	227	299	0.5469	0.7204
10740	887,077	Albuquerque, NM	485	85	0.5467	0.0958
31140	1,235,708	Louisville/Jefferson County, KY-IN	670	795	0.5422	0.6434
38900	2,226,009	Portland-Vancouver-Hillsboro, OR-WA	1,202	1,383	0.54	0.6213
23540	264,275	Gainesville, FL	142	134	0.5373	0.507
17460	2,077,240	Cleveland-Elyria, OH	1,111	1,257	0.5348	0.6051
42660	3,439,809	Seattle-Tacoma-Bellevue, WA	1,831	1,285	0.5323	0.3736
18140	1,901,974	Columbus, OH	1,012	875	0.5321	0.46
28140	2,009,342	Kansas City, MO-KS	1,069	551	0.532	0.2742
10420	703,200	Akron, OH	374	211	0.5319	0.3001
12100	274,549	Atlantic City-Hammonton, NJ	145	169	0.5281	0.6156
40900	2,149,127	Sacramento-Roseville-Folsom, CA	1,133	1,445	0.5272	0.6724
21140	197,559	Elkhart-Goshen, IN	103	57	0.5214	0.2885
17820	645,613	Colorado Springs, CO	331	208	0.5127	0.3222
24660	723,801	Greensboro-High Point, NC	366	332	0.5057	0.4587
16820	218,705	Charlottesville, VA	109	236	0.4984	1.0791
17140	2,114,580	Cincinnati, OH-KY-IN	1,036	863	0.4899	0.4081
38860	514,098	Portland-South Portland, ME	249	298	0.4843	0.5797
44140	621,570	Springfield, MA	299	300	0.481	0.4826
35980	274,055	Norwich-New London, CT	131	86	0.478	0.3138
19660	590,289	Deltona-Daytona Beach-Ormond Beach, FL	282	333	0.4777	0.5641
16580	231,891	Champaign-Urbana, IL	109	65	0.47	0.2803
15500	151,131	Burlington, NC	71	23	0.4698	0.1522
33260	141,671	Midland, TX	66	225	0.4659	1.5882
42020	269,637	San Luis Obispo-Paso Robles, CA	124	222	0.4599	0.8233
18880	235,865	Crestview-Fort Walton Beach-Destin, FL	107	159	0.4536	0.6741
14540	158,599	Bowling Green, KY	71	82	0.4477	0.517
41180	2,787,701	St. Louis, MO-IL	1,242	239	0.4455	0.0857
48900	254,884	Wilmington, NC	110	91	0.4316	0.357
14260	616,561	Boise City, ID	256	278	0.4152	0.4509
35380	1,189,866	New Orleans-Metairie, LA	493	401	0.4143	0.337
22660	299,630	Fort Collins, CO	124	110	0.4138	0.3671
11700	424,858	Asheville, NC	161	110	0.379	0.2589
39900	425,417	Reno, NV	160	177	0.3761	0.4161
40580	152,392	Rocky Mount, NC	56	42	0.3675	0.2756
26900	1,887,877	Indianapolis-Carmel-Anderson, IN	693	713	0.3671	0.3777
15260	112,370	Brunswick, GA	41	33	0.3649	0.2937
32820	1,324,829	Memphis, TN-MS-AR	477	904	0.36	0.6824

16860	528,143	Chattanooga, TN-GA	189	286	0.3579	0.5415
27980	154,924	Kahului-Wailuku-Lahaina, HI	53	91	0.3421	0.5874
46520	953,207	Urban Honolulu, HI	326	537	0.342	0.5634
42340	347,611	Savannah, GA	116	91	0.3337	0.2618
29180	466,750	Lafayette, LA	155	90	0.3321	0.1928
20020	145,639	Dothan, AL	48	48	0.3296	0.3296
26140	141,236	Homosassa Springs, FL	46	53	0.3257	0.3753
45220	367,413	Tallahassee, FL	117	141	0.3184	0.3838
41700	2,142,508	San Antonio-New Braunfels, TX	682	750	0.3183	0.3501
40140	4,224,851	Riverside-San Bernardino-Ontario, CA	1,342	1,527	0.3176	0.3614
17860	162,642	Columbia, MO	51	6	0.3136	0.0369
39460	159,978	Punta Gorda, FL	50	35	0.3125	0.2188
17660	138,494	Coeur d'Alene, ID	43	38	0.3105	0.2744
21780	311,552	Evansville, IN-KY	96	75	0.3081	0.2407
33860	374,536	Montgomery, AL	115	314	0.307	0.8384
13460	157,733	Bend, OR	47	87	0.298	0.5516
19430	799,232	Dayton-Kettering, OH	237	164	0.2965	0.2052
29820	1,951,269	Las Vegas-Henderson-Paradise, NV	560	725	0.287	0.3716
45780	610,001	Toledo, OH	175	242	0.2869	0.3967
17980	294,865	Columbus, GA-AL	84	106	0.2849	0.3595
29460	602,095	Lakeland-Winter Haven, FL	171	175	0.284	0.2907
37860	448,991	Pensacola-Ferry Pass-Brent, FL	127	110	0.2829	0.245
33460	3,348,859	Minneapolis-St. Paul-Bloomington, MN-WI	942	3,250	0.2813	0.9705
33660	412,992	Mobile, AL	114	108	0.276	0.2615
36100	331,298	Ocala, FL	91	86	0.2747	0.2596
33700	514,453	Modesto, CA	141	118	0.2741	0.2294
19300	182,265	Daphne-Fairhope-Foley, AL	49	50	0.2688	0.2743
41100	138,115	St. George, UT	37	31	0.2679	0.2245
27620	149,807	Jefferson City, MO	40	38	0.267	0.2537
37460	184,715	Panama City, FL	49	54	0.2653	0.2923
28940	837,571	Knoxville, TN	215	261	0.2567	0.3116
49180	640,595	Winston-Salem, NC	164	125	0.256	0.1951
40060	1,208,101	Richmond, VA	308	570	0.2549	0.4718
40420	349,431	Rockford, IL	89	74	0.2547	0.2118
27140	567,122	Jackson, MS	144	142	0.2539	0.2504
43780	319,224	South Bend-Mishawaka, IN-MI	81	59	0.2537	0.1848
44100	210,170	Springfield, IL	53	47	0.2522	0.2236
48620	630,919	Wichita, KS	151	120	0.2393	0.1902
27740	198,716	Johnson City, TN	47	42	0.2365	0.2114
23580	179,684	Gainesville, GA	42	37	0.2337	0.2059
36260	597,159	Ogden-Clearfield, UT	138	145	0.2311	0.2428
45820	233,870	Topeka, KS	54	308	0.2309	1.317
12620	153,923	Bangor, ME	35	28	0.2274	0.1819
37900	379,186	Peoria, IL	86	55	0.2268	0.145
39300	1,600,852	Providence-Warwick, RI-MA	361	405	0.2255	0.253
29200	201,789	Lafayette-West Lafayette, IN	45	40	0.223	0.1982
32780	203,206	Medford, OR	45	48	0.2215	0.2362
29740	209,233	Las Cruces, NM	46	12	0.2199	0.0574
38060	4,192,887	Phoenix-Mesa-Chandler, AZ	920	882	0.2194	0.2104
12940	802,484	Baton Rouge, LA	175	163	0.2181	0.2031
29940	110,826	Lawrence, KS	24	11	0.2166	0.0993
30340	107,702	Lewiston-Auburn, ME	23	20	0.2136	0.1857
14020	159,549	Bloomington, IN	34	40	0.2131	0.2507
21660	351,715	Eugene-Springfield, OR	72	104	0.2047	0.2957
40220	308,707	Roanoke, VA	63	33	0.2041	0.1069
14740	251,133	Bremerton-Silverdale-Port Orchard, WA	51	31	0.2031	0.1234
44420	118,502	Staunton, VA	24	12	0.2025	0.1013
23060	416,257	Fort Wayne, IN	84	122	0.2018	0.2931
22180	366,383	Fayetteville, NC	73	54	0.1992	0.1474
31420	232,293	Macon-Bibb County, GA	46	42	0.198	0.1808
12020	192,541	Athens-Clarke County, GA	37	42	0.1922	0.2181
22140	130,044	Farmington, NM	25	3	0.1922	0.0231
10540	116,672	Albany-Lebanon, OR	22	27	0.1886	0.2314
10580	870,716	Albany-Schenectady-Troy, NY	160	266	0.1838	0.3055
44180	436,712	Springfield, MO	80	16	0.1832	0.0366
33740	176,441	Monroe, LA	32	20	0.1814	0.1134
15940	404,422	Canton-Massillon, OH	72	50	0.178	0.1236
19460	153,829	Decatur, AL	27	25	0.1755	0.1625
31340	252,634	Lynchburg, VA	44	23	0.1742	0.091
23420	930,450	Fresno, CA	159	274	0.1709	0.2945
13980	178,237	Blacksburg-Christiansburg, VA	30	10	0.1683	0.0561
28700	309,544	Kingsport-Bristol, TN-VA	52	53	0.168	0.1712
25860	365,497	Hickory-Lenoir-Morganton, NC	60	69	0.1642	0.1888
31740	92,719	Manhattan, KS	15	2	0.1618	0.0216
12540	839,631	Bakersfield, CA	134	140	0.1596	0.1667
44060	527,753	Spokane-Spokane Valley, WA	83	57	0.1573	0.108
47260	1,676,822	Virginia Beach-Norfolk-Newport News, VA-NC	260	349	0.1551	0.2081
44700	685,306	Stockton, CA	106	170	0.1547	0.2481
43340	439,811	Shreveport-Bossier City, LA	67	76	0.1523	0.1728
25060	370,702	Gulfport-Biloxi, MS	56	37	0.1511	0.0998
24540	252,825	Greeley, CO	38	31	0.1503	0.1226
49660	565,773	Youngstown-Warren-Boardman, OH-PA	80	47	0.1414	0.0831
25620	142,842	Hattiesburg, MS	20	30	0.14	0.21
24420	82,713	Grants Pass, OR	11	14	0.133	0.1693
12260	564,873	Augusta-Richmond County, GA-SC	75	42	0.1328	0.0744
46220	230,162	Tuscaloosa, AL	30	59	0.1303	0.2563
26380	208,178	Houma-Thibodaux, LA	27	19	0.1297	0.0913

26580	364,908	Huntington-Ashland, WV-KY-OH	45	30	0.1233	0.0822
47380	252,772	Waco, TX	31	39	0.1226	0.1543
15380	1,135,509	Buffalo-Cheektowaga, NY	130	46	0.1145	0.0405
18580	428,185	Corpus Christi, TX	49	49	0.1144	0.1144
21340	804,123	El Paso, TX	92	226	0.1144	0.2811
46060	980,263	Tucson, AZ	112	92	0.1143	0.0939
19340	379,690	Davenport-Moline-Rock Island, IA-IL	43	42	0.1133	0.1106
41420	390,738	Salem, OR	38	56	0.0973	0.1433
13780	251,725	Binghamton, NY	21	3	0.0834	0.0119
13140	403,190	Beaumont-Port Arthur, TX	29	25	0.0719	0.062
11260	380,821	Anchorage, AK	23	79	0.0604	0.2074
40380	1,079,671	Rochester, NY	59	42	0.0546	0.0389
45060	662,577	Syracuse, NY	26	26	0.0392	0.0392
27060	101,564	Ithaca, NY	3	6	0.0295	0.0591

Table A4: How Does the Estimated Quality of Movers that Leave and Arrive to a City Correlate?

	Avg. In Mover Quality		
	(1)	(2)	(3)
Avg. Out Mover Quality	0.023 (0.064)	-0.040 (0.072)	0.014 (0.056)
Avg. Quality of Local Firms		0.570*** (0.170)	0.823*** (0.103)
Weights			Total Movers
Observations	171	171	171
R ²	0.001	0.093	0.570

OLS regression. Average quality estimated by replicating the measure of Guzman and Stern (2020) in the data. Specifically, for all non-movers born before 2012, we run a logit model with a binary measure of equity events as the dependent variable, and observables for whether a firm, close to founding and in its birth location, is a corporation, has a short name, is eponymous, has a patent, has a trademark, has both a patent and a trademark, and five industry characteristics based on firm name. Predictions from this model report an out of sample ROC score of 0.80. Estimated quality is the predicted out of sample probability of this model. We average this value for all movers in and out of a city, and firms born in a city that do not move. Robust standard errors in parenthesis. Column (3) is weighted by the total movers in or out of each city. Significance denoted as *p<0.1; **p<0.05; ***p<0.01

Data Appendix To: Entrepreneurial Migration

Kevin Bryan
(University of Toronto)

Jorge Guzman
(Columbia University)

1 Introduction

This appendix overviews the construction and development of the data in our paper *Entrepreneurial Migration*. The appendix is divided into four sections. First, we cover the conceptual goal and need for measuring entrepreneurial migration. Then, we outline the key challenges in doing so, particularly around firm heterogeneity, defining migration, and observability. Then, we explain the data — business registration records for Delaware registered companies — and the overall approach to constructing our dataset. We also review the key summary statistics of the full set of firms. Finally, we compare our data to other potential datasets. Abridged fragments of this appendix are also included in the main text.

2 Why Measure Migration of High Growth Startups?

The importance of understanding the role of location on startup performance has been of interest at least since Marshall (Marshall, 1890; Jacobs, 1970; Saxenian, 1994; Glaeser, Kerr, & Kerr, 2015). A growing literature documents a number of localized economic benefits for regions that have more startups, the most important one being economic growth (Glaeser, Kerr, & Ponzetto, 2010). Over the last decade, an important formalization of this relationship has emphasized that it is one group of startups in particular — high growth

startups — that account for the bulk of this economic impact (Schoar, 2010; Guzman & Stern, in press). High growth startups are firms that have a disproportionate likelihood of growth. In particular, a number of studies have documented that this growth intent is reflected in founding choices entrepreneurs take in the early stages of their business activities (Guzman & Stern, in press).

In direct contrast to the importance of the location where firms locate is the possibility of migration. While most startups are born and develop in the city where their founders lived prior to founding (Michelacci & Silva, 2007), this pattern is not universal. Anecdotes abound of good entrepreneurs who chose to start a company in one location only to see it grow in a different one. For example, while Marc Andreessen had all the initial ideas and training for what would become Netscape at the University of Illinois Urbana-Champaign, he moved to California to build the company itself. Similarly, Bill Gates and Paul Allen wrote the original Microsoft programs while Gates was a student at Harvard, but they eventually grew the company in the Seattle area. The impact of these entrepreneurial migrations on their destination regions has been substantial. A series of policies has emerged to motivate high talent entrepreneurs to move to a region with the goal of replicating some version of this story, the most notable of these being Startup Chile. Other policies, most notably in Israel, instead encourage entrepreneurs to ‘move out’ of their home region to a richer destination with the goal that the spillovers from future growth benefit back into the region (Conti & Guzman, 2019).

Yet, whether migration of high growth startups actually happens, and what are the characteristics that drive it, appear so far unexplored.

Understanding the economic phenomenon of entrepreneurial migration poses a number of both conceptual and measurement challenges. Migration has been studied substantially in economic theory (e.g. Roback) as a choice problem over some maximization function for either people or firms. Absent principal-agent issues, this maximization should be over the weighted utility of the equity-holders of the firm. Yet, because entrepreneurs also tend to be the managers, the maximization cannot simply be done on the role of location on increasing firm value, but also on the utility costs for managers to relocate to one of these

regions, independent of the startup. For example, relocating might require being away from loved ones, losing an additional personal income source (e.g. the income of a spouse), or simply living in a location that is not personally desirable. Furthermore, these same personal connections also constitute valuable local relationships, that in and of themselves are likely to impact firm performance.

To date, a series of studies has emerged understand the differences between personal connections and locational benefits as drivers to startup firm performance (Dahl & Sorenson, 2012; Michelacci & Silva, 2007; Guzman, 2018), as well as how changes in the 'appeal' of a city influence would-be migrants on their choice of hiring a manager or moving themselves (Kulchina, 2014). However, a systematic measurement of entrepreneurial migration for high growth startups does not yet exist, leaving many critical questions unanswered.

3 The Difficulty in Measuring Startup Migration

Measuring entrepreneurial migration itself represents a few unique challenges, including accounting for firm quality, observing firms in their original location, and observing the migration of the firms in a timely fashion. We review each in turn.

Accounting for firm quality in migration is particularly important. One reason is the growing sense of importance assessed in the literature to the significant heterogeneity in firm potential (Schoar, 2010; Guzman & Stern, in press) – with a few 'high growth' firms accounting for the majority of the economic impact of entrepreneurship. Understanding the migration patterns of all firms might explain little about economic growth, while finding the few firms that do have the potential to grow might be much more informative. A second, equally important, reason is that the motivations for migration, or the behaviors that lead to them, might be different across the entrepreneurship of high growth and non high growth startups. Recent evidence finds ample variation on the personality of high growth entrepreneurs versus other types of actors (Kerr, Kerr, & Xu, 2017), and studies on the motivations of these shows that it is not only profit or productivity that defines their choices (Guzman, Oh, & Sen, 2020). In short, a clear focus on measurement of high

growth startups is critical to understand the phenomenon of entrepreneurial migration and its performance.

The remaining concerns reflect challenges in the observability of entrepreneurial migrations. Because some founders move before starting a company, while other migrants might become entrepreneurs only years after arriving in an entrepreneurial region (Saxenian, 2007), there is no obvious breakpoint on which to define a migration as ‘entrepreneurial’. A different, narrower approach, and the one we focus on in this paper, focuses on simply studying the migration of newly born startups. The unique advantage of using this definition is that it circumvents vexing questions about how location influences the choice of entrepreneurial entry. That is, if individuals migrate before becoming entrepreneurs, would they have been entrepreneurs before migrating?

Finally, there remains a question of how to observe the changes in the location of firms. That is, restricting ‘entrepreneurial migration’ to mean a firm that moves its headquarters to a new location, the problem involves defining “firm”, “headquarters”, and tracking these moves in a consistent way. We take advantage of institutional details in the United States that allow this tracking.

4 Data

Our analysis is focused on the founding and geographic reallocation of companies registered under Delaware jurisdiction. These are not companies headquartered in Delaware — they are headquartered across the United States. Instead, being under Delaware jurisdiction reflects the fact that when a firm is founded it has the freedom to choose where to register.¹ This choice of jurisdiction is consequential to a large number of corporate legal aspects of the firm, including labor disputes, shareholder disputes, and the legality and enforceability of certain contracts. Since the early twentieth century, two broad choices of jurisdiction have emerged for new U.S. firms.

¹This feature of multiple jurisdictions appears to be an unusual feature of the United States. In most other countries, corporate law is overwhelmingly similar across all regions of the country.

Most startups (about 96%) initially register under only the local jurisdiction of their own state. There are several benefits to registering in the local jurisdiction, including a simplicity in translating between corporate law and the local law, and the need to pay for only one registration. In general, being in the local jurisdiction is simply cheaper.

A few companies (most of the remainder 4%), however, choose instead to register under Delaware jurisdiction and then operate as a foreign (out of state) company in the state in which they are headquartered. This process is more expensive, as it requires more legal work to maintain both registrations, and the firms need to pay fees to both states. However, it also creates certain benefits that accrue particularly well to entrepreneurs that intend to scale the company. First, corporate law is mostly case-based in the United States, and Delaware is the state with the largest canon of corporate law. This means that precedent on the enforceability of different clauses and contracts has been tested and developed in detail. Venture capitalists, for example, are usually reticent to extend contracts to firms in other jurisdictions due to the uncertainty of knowing whether and how a contract would hold. Second, Delaware Corporate Law is commonly taught in law schools nationwide. Finally, Delaware has a reputation for fairness in dealing with corporate disputes, through its specialized Court of the Chancery. Together, these benefits have become significant for many firms in the United States, and are particularly valuable for those firms that intend to be large. The additional costs of Delaware registration create a separating equilibrium of sorts: firms with high growth intention choose Delaware, while the rest choose the local law (Catalini, Guzman, & Stern, 2019). Accordingly, while Delaware represents less than 0.5% of the U.S. population, over half of all U.S. publicly listed firms are registered here. In empirical estimates, firms registered in Delaware at founding are over 45 times more likely to achieve an equity growth outcome (such as an IPO or acquisition) (Guzman & Stern, in press).

To track the migration of Delaware firms in their location choices, we take advantage of unique institutional rules in state-level corporate laws, requiring firms to register in every state in which they engage in meaningful business activity.² These registrations are re-

²Broadly, this occurs when a firm has hired employees in a state, opened a bank account, or is renting an office.

quired to use exactly the same official firm name, down to the comma, in each state where they do business. Because firms register in a state only at the time of entering the state, we can use the registration date to assess when a firm expands location to another state. In most cases, this is a subsidiary expansion while the headquarters of the company remain in the home location. Yet, sometimes, it will represent (or will eventually become) an entrepreneurial migration — i.e. the relocation of the company headquarters.

Differentiating between these two modes of expansion is difficult as it would require a firm to state separately the location of the principal office and the location of the state-based office. Through a manual check of the information in each state, we identified 35 U.S. states, and the District of Columbia, in which the process of registration distinctly requires firms to separately document the local state office and the principal office through one of two modes: either requiring the primary corporate address explicitly in the registration form, or by requesting the address of the president, CEO, or main manager of the firm. In the latter case, we assume that if the majority of officers live in the same MSA, then the corporate headquarters is located in that MSA. These 36 jurisdictions form the basis of our analysis.³

We obtained data on all the Delaware jurisdiction firms registered between 1988 and 2014 in each of these states through the Startup Cartography Project (Andrews, Fazio, Guzman, Liu, & Stern, 2019). The Startup Cartography Project (SCP) is a project focused on the measurement of firm formation through business registration across time and location. The data included the name of each company, the registered address of the principal office, and the date in which it registered in each state. We also obtained all observables used in the SCP to measure entrepreneurial quality – an estimate of the founding potential of companies based on the predicted probability of growth based on founding characteristics.

Using this information, we matched the companies across each state in our sample. To

³Our states are Alaska, Alabama, Arizona, California, Colorado, Connecticut, DC, Florida, Georgia, Hawaii, Idaho, Illinois, Indiana, Kansas, Kentucky, Louisiana, Massachusetts, Maine, Minnesota, Missouri, Mississippi, North Carolina, New Hampshire, New Jersey, Nevada, New York, Ohio, Oregon, Rhode Island, Tennessee, Texas, Utah, Virginia, Vermont, Washington, and West Virginia.

do this, we tracked the initial state registration date of each firm in each state, as well as the registered zip code (either of the “primary” company address, or the broader MSA in the case of states where director addresses were used). Using this data, we operationalize a measure of migration through the following algorithm:

1. The first state in which the firm is registered is the founding state.
2. If a firm changes its principal office to another state, and the destination MSA does not include the source state, we consider this a migration.
3. The date in which it first registers in the destination state is the migration date.

This allows us to track well the relocation of startup companies across state-lines. In our main analysis, a startup migration is a firm that moves within five years of the first time they appear in our data. We drop all moves within 3 months of the initial founding date as these tend to conflate moves with firms who register in many states on founding. For instance, a restaurant chain that spins out one of its brands as an independent firm will be registered in many states nearly simultaneously. The fact that one state processes the registration a few days before another does not mean that the firm was “founded” in the earlier state.

5 Drawbacks and Risks of Measurement Approach

Our approach does come with several drawbacks and potential risks. We review each of the main ones in turn.

The timing of migration. In the process of migration, timing is important. Our data does not allow us to know precisely the date a company changes the official main location for a company. Indeed, this “precise date” is not particularly well-defined. A Seattle-based startup may open an office in Phoenix in 2007, slowly move various corporate tasks to that office in 2008, then begin referring to Phoenix as its “headquarters” publicly in 2009. However, for the analyst it is not conceptually obvious when the headquarter “move” began. We therefore define the date of a move as the first date a firm registers business in

any state where it eventually refers to that state as housing its “principal address”.

We believe this is a relatively minor concern because it does not affect *who* we code as migrants, nor *where* do they move to, but only *when* they move. The timing of migration itself is not a main area of analysis in our paper.

Relocation within states. The strength of our data is in identifying migration across state lines to different MSAs. Our data, however, does not allow us to track migrations within the same state such as moving from San Diego, CA to San Francisco, CA or from Rochester, NY to New York City. Although restricting to cross-state migrations limits the total number of HQ moves, it does not bias the results of our utility-based approach. Recall that the utility-based approach depends on the relative number of moves between cities, and omitting within-state moves means the omitted moves are bi-directional for any city pair. Note also that when tracking MSA moves, we also drop moves if the firm moves from one MSA to a different state which also makes up the origin MSA. For instance, the New York City MSA includes zip codes in New Jersey, so a New York City firm that moved to New Jersey will not be counted as a cross-state move in our data. This is due to issues with interpolating origin MSAs when only the state of origin can be observed, as noted below. Again, this omission does not bias our results.

What (and who) moves? Another limitation of our data is that it does not allow us to go into the organizational structure of each migration beyond the relocation of headquarters. Naturally, some firms will not move fully and might leave someone in the original location, or might choose other work arrangements. Future datasets would do well to improve upon this margin.

Definition of a startup. We define startup, as discussed, to mean a new business entity. Spinouts and subsidiaries of existing firms, which may be quite large at “founding”, therefore count as startups. Hand-investigation of the data suggests that the vast majority of data points are “true startups”, meaning small, de novo firms. That said, utility estimates for some cities are affected by this distinction. For example, Peoria, Illinois is the highest utility small city, based largely on having 12 startups move in while only 2 move out. Many of these 12 moves are the result of agricultural acquisitions, whereby a novel

corporate entity was created to help facilitate the sale, and the headquarters was then integrated into Peoria the next year. Moves of this type are, however, quite rare in the data at large.

References

- Andrews, R., Fazio, C., Guzman, J., Liu, Y., & Stern, S. (2019). The startup cartography project: Measuring and mapping entrepreneurial ecosystems. *Working Paper*.
- Catalini, C., Guzman, J., & Stern, S. (2019). *Passive versus active growth: Evidence from founder choices and venture capital investment* (Tech. Rep.). National Bureau of Economic Research.
- Conti, A., & Guzman, J. (2019). What is the us comparative advantage in entrepreneurship? evidence from israeli migration to the united states. *Working Paper*.
- Dahl, M. S., & Sorenson, O. (2012). Home sweet home: Entrepreneurs' location choices and the performance of their ventures. *Management science*, 58(6), 1059–1071.
- Glaeser, E. L., Kerr, S. P., & Kerr, W. R. (2015). Entrepreneurship and urban growth: An empirical assessment with historical mines. *Review of Economics and Statistics*, 97(2), 498–520.
- Glaeser, E. L., Kerr, W. R., & Ponzetto, G. A. (2010). Clusters of entrepreneurship. *Journal of urban economics*, 67(1), 150–168.
- Guzman, J. (2018). Go west young firm: The value of entrepreneurial migration for startups and founders. *Working Paper*. Working Paper.
- Guzman, J., Oh, J. J., & Sen, A. (2020). What motivates innovative entrepreneurs? evidence from a global field experiment. *Management Science*, 66(10), 4808–4819.
- Guzman, J., & Stern, S. (in press). The state of american entrepreneurship: New estimates of the quantity and quality of entrepreneurship for 32 us states, 1988-2014. *American Economic Journal: Economic Policy*.
- Jacobs, J. (1970). *The economy of cities*. Penguin Random House.
- Kerr, S. P., Kerr, W. R., & Xu, T. (2017). Personality traits of entrepreneurs: A review of recent literature.
- Kulchina, E. (2014). Media coverage and location choice. *Strategic Management Journal*, 35(4), 596–605.
- Marshall, A. (1890). *The principles of economics*. McMaster University Archive for the History of Economic Thought.
- Michelacci, C., & Silva, O. (2007). Why so many local entrepreneurs? *The Review of Economics and Statistics*, 89(4), 615–633.
- Saxenian, A. (1994). *Regional advantage: Culture and competition in silicon valley and route 128*. Harvard University Press.
- Saxenian, A. (2007). *The new argonauts: Regional advantage in a global economy*. Harvard University Press.
- Schoar, A. (2010). The divide between subsistence and transformational entrepreneurship. *Innovation policy and the economy*, 10(1), 57–81.