Review of Economics and Statistics Just Accepted MS.
https://doi.org/10.1162/rest_a_01381

Entrepreneurial Migration*

Kevin A. Bryan† Jorge Guzman‡

September 25, 2023

Abstract

We track the movement of high-potential startups using cross-state business registrations and estimate the utility of cities to moving startups using a revealed preference approach. 6.6% of these startups move across state borders during their first five years. Startup hubs like Silicon Valley and Boston tend to *lose* startups to other cities. Our findings show that startups prefer traditional hubs when they move soon after being founded, but later prefer cities with lower taxes. This pattern is not due to vertical sorting or industrial specialization. ECONLIT CODES: L26, R12, R13

*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 for helpful comments three anonymous referees, and Thomas Astebro, Pierre Azoulay, 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 University of Cape Town, 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 and NBER. jag2367@gsb.columbia.edu.

Review of Economics and Statistics Just Accepted MS.
https://doi.org/10.1162/rest_a_01381

1 Introduction

Most high-growth startups are born in a small set of cities (Guzman and Stern, 2020). These "startup hubs" often have resources that are difficult for other cities to replicate, such as major technical universities, keystone firms that spin out founders, or a venture finance ecosystem. That said, in addition to *birthing* startups, cities can also *attract* young, fast-growing firms to move from their initial location. While most startups remain near where they were born (Dahl and Sorenson, 2012; Michelacci and Silva, 2007; Guzman, 2023), some prominent firms, including Microsoft, Facebook, and Slack, moved when they were young. This raises several questions: how common is entrepreneurial migration? What cities do startups prefer when they move? And what leads startups 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

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

Review of Economics and Statistics Just Accepted MS.
https://doi.org/10.1162/rest_a_01381

> 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 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 pool is sufficiently deep.

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

[2]https://www.kansascity.com/news/local/article326075.html

[3]https://technical.ly/dc/2018/06/20/vbrick-20-million/

[4]The 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.

Review of Economics and Statistics Just Accepted MS.
https://doi.org/10.1162/rest_a_01381

(2023) 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 about as often as working-age adults. Second, the rich do not get richer: more startups move *out* of major hubs like Boston, San Francisco, and Silicon Valley than move *to* these hubs. Third, very new startups and slightly more advanced ones do not value cities equally when they move. While very young startups move to traditional hubs, startups between their third and fifth years after firm formation are much more likely to move to low-tax jurisdictions. This final pattern is seen most strongly in the highest-quality startups as measured by growth intention at birth.

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, and further can identify the relative utility of startups who move soon after being founded from those who move later. 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

Review of Economics and Statistics Just Accepted MS.
https://doi.org/10.1162/rest_a_01381

will suggest that one city is "better" for entrepreneurs than another if it attracts firms from other good cities, and loses few firms. Further, the model suggests both that the highest-quality firms of any vintage are more likely to move and that the probability a startup moves falls in the age of the firm.

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]

To explain this pattern, note that startups are unusual in that they face a fundamental

[5] As 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.

Review of Economics and Statistics Just Accepted MS.
https://doi.org/10.1162/rest_a_01381

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. Of course, the majority of firms do not move at all from where they begin operations, due to switching costs.

In line with this tension, we establish the following empirical regularities. When young firms move, they go to startup hubs, meaning cities with an existing agglomeration of startups. This pattern is particularly strong for the highest-quality young firms. As the firms age, their relocation decisions tilt toward cities that are more business-friendly. That is, for the highest-quality startups, we observe patterns consistent with a *"nursery cities"* model à la 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. Lower-quality startups weigh amenities differently, and on the margin are more likely to move to cities that provide high utility for the founder via factors like weather or overall amenities as in Albouy (2016).

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, 2023; Conti and Guzman, 2023), but this work has remained purely focused on identifying the impact of moving on individual startup performance. Relative to this

Review of Economics and Statistics Just Accepted MS.
https://doi.org/10.1162/rest_a_01381

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.

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. 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 their founding: in both cases, the relative utility we estimate is the most appropriate theoretical construct.

Review of Economics and Statistics Just Accepted MS.
https://doi.org/10.1162/rest_a_01381

2 Three Facts about Entrepreneurial Migration

Before deriving a theory of entrepreneurial migration and explaining our data in detail, we begin by documenting three facts on the migration of high-growth startups across the United States. 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. (2022) and references therein).

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

This figure is our baseline estimated migration rate for startups 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.[6] We then use the predicted value of this regression to estimate the number of movers leaving each of the states in our data 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).[7]

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

[6]In Appendix Table A9, we show that conditional on making a cross-state move, distance between origin and destination city plays only a tiny role in move rates.

[7]While our dataset only includes U.S. firms registered in Delaware at birth, Braun and Weik (2021) examine HQ moves of venture-backed European startups and find that a minimum of 3.5% of all European venture-backed startups since 2000 moved to the US. Incredibly, every European company in their sample has net outmigration.

Review of Economics and Statistics Just Accepted MS.
https://doi.org/10.1162/rest_a_01381

Figure 1 shows this graphically, plotting startup births and net startup moves per capita in large CBSAs. In both pictures, high per capita figures are represented by blue and low per capita figures by white. 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.

Fact 3. *There is no correlation overall between startup hubs and net startup moves. Instead, very young startups are more likely to move to hubs, and older startups are more likely to move to business-friendly locations.*

Figure 2 plots the net migration ratio — the number of arrivals divided by the number of departures within five years of a company's founding — against the number of startups founded in each city per capita. There is in general no correlation between the two. However, this null result hides an important lifecycle effect, which we will demonstrate in Section 5: cities with many startups per capita see net inmigration of startups less than two years old, but net outmigration of startups between their third and fifth year after birth. Cities with few startups born there see the opposite pattern.

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.

3 A Revealed Preference Model of Startup Migration

Rather than fitting a hedonic gravity model to estimate the utility of cities to startups, which would require prespecifying covariates which determine that utility, we construct a rank of cities based purely on revealed preference. In particular, we will assume that the spatial

Review of Economics and Statistics Just Accepted MS.
https://doi.org/10.1162/rest_a_01381

distribution of startups at birth is driven by idiosyncratic firm-level factors plus common city-level fixed effects. We then assume that startups in any given period consider moving when the expected utility of doing so exceeds the cost of investigating where to move. If they pay this cost, 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).

There are benefits and costs of a revealed preference rather than a hedonic approach. A hedonic approach to firm migration requires the analyst to prespecify the firm- and city-level variables she expects to matter to the firm's migration choice. It is not at all obvious ex-ante what these should be. On the other hand, incorporating firm-level heterogeneity is straightforward in a hedonic model. Our revealed preference approach gains the ability to rank-order cities in a formally identified way at the cost of ruling out heterogeneous yet correlated preferences across firms over cities. We discuss theoretically in this section, and empirically in Section 5, how this limitation affects the interpretation of our results.

3.1 Model Assumptions

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

Review of Economics and Statistics Just Accepted MS.
https://doi.org/10.1162/rest_a_01381

wages and firm-to-firm voluntary transitions. The model requires three assumptions, one about how startups are born, one about how they decide whether to investigate moving, and one about what costly information they receive when they consider moving.

Assumption 1. *Assume there are N potential entrepreneurs in society, and J cities. 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 theory, 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. *In future periods t, startups consider whether to move. By paying a firm- and time-specific cost C_{it} drawn from a distribution F_t, startups will receive another utility draw from each city equal to $V_{jt} + \epsilon_{ijt}$. As before, ϵ are draws from a mean-zero i.i.d. extreme value type 1 distribution with scale 1.*

We will call V_j the "common utility" component of a city's utility to a startup, and ϵ_{ij} the idiosyncratic component. This assumption implies that firms can, in each period, acquire information about the value of moving to a different city at a cost C_{it}, which we interpret as a cost per unit of expected future profitability.[8] Firms consider moving as long as the

[8]Note that utilities $V_{jt} + \epsilon_{ijt}$ are normalized and not, for example, scaled by revenue or expected future profit.

Review of Economics and Statistics Just Accepted MS.
https://doi.org/10.1162/rest_a_01381

expected payoff exceeds the cost of this search.

Assumption 3. *Before deciding whether to pay the cost of acquiring information about the value of potential moves, startups hold the prior that the common utility of all cities in the following period is drawn from $V_{jt} = v + \gamma_{jt}$, where γ_{jt} are i.i.d., normally distributed with mean zero and variance σ.*

The final assumption says that firms hold the uninformative prior each period that all cities are ex-ante equally likely to be good for startups of their particular birth-year, and that the common utility of cities will be normally distributed. A higher σ means that firms believe the common utility of cities will be more variable. Note that there is no persistence in beliefs across periods: a very young startup that investigated cities and learned that Louisville was a high-utility city for them at the time would nonetheless have only an uninformative prior about whether Louisville or New Orleans would be more promising five years after the firm was founded.

Putting these assumptions together, the model says that potential entrepreneurs are either born or else stay in non-entrepreneurial employment. Each period after birth, these startups can move cities if they like, but they only learn the utility benefits of moving by paying a cost. The firms that pay for this investigation 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).

The fact that the idiosyncratic component is uncorrelated across cities and time 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 largely pure industrial sorting, this model is inappropriate. In that case, 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.[9] However, if the industrial *diversity* of a city, or the level of industrial specialization, or the amenity

[9]Note that it is possible to perform the algorithm described in this section industry-by-

Review of Economics and Statistics Just Accepted MS.
https://doi.org/10.1162/rest_a_01381

value of cities is what drives firm moves, and firms merely differ in the importance they place idiosyncratically on those features, the assumption holds. Note also that our model identifies utility solely from revealed preference of movers; in our city-firm matching process, there is no equivalent of wages in the firm-worker matching process of Abowd et al. (1999) and the literature that followed. These models permit heterogeneity in preferences at the level of the mover beyond the fact that origins differ in their propensity to generate moves, but use wages to close the model.

3.2 Deriving Utility from Revealed Preference

We now show how to analytically extract the vector of *relative utilities* $\bar{V_t} = V_t - B$. 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. If $\bar{V_{jt}}$ is positive, startups get higher utility on average in period t 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, and how that attractiveness varies at different times in the startup lifecycle.

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

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 use this heterogeneity in our primary results, but will discuss robustness to limiting the data to IT and Health industry subsamples in Section 5.

Review of Economics and Statistics Just Accepted MS.
https://doi.org/10.1162/rest_a_01381

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 $\bar{V}_t$, beginning with the decision to consider moving. A firm will pay the cost of moving C_{it} if the expected payoff to moving is sufficiently high. Note that since beliefs about the utility of cities are not persistent across periods, we can analyze this decision myopically. In particular, a firm i will move if

$$\mathbb{E}[\max_j(V_{jt} + \epsilon_{ijt}) - V_{ct} - \epsilon_{ict}] \geq C_{it}$$

where c is the current city the firm is considering leaving. That is, the expected utility of the best draw they receive needs to be at least C_{it} higher than their expected utility from remaining in the current city c.

Using the prior that $V_{jt} = v + \gamma_j$, the firm will consider moving if and only if

$$\mathbb{E}[\max_j(v + \gamma_j + \epsilon_{ijt}) - v - \gamma_c - \epsilon_{ict}] = \mathbb{E}[\max_j(\gamma_j - \gamma_c + \epsilon_{ijt} - \epsilon_{ict})] \geq C_{it}$$

Since the difference of two standard Gumbel distributions is a standard logistic, and the difference of the two mean-zero normal distributions is a mean-zero normal with standard deviation $\sqrt{1+\sigma^2}$, we have that the firm will consider moving if and only if

$$\mathbb{E}[\Omega(\sigma)] \geq C_{it}$$

where $\Omega(\sigma)$ is the maximum of J symmetric i.i.d. random variables whose distributions are the sum of a standard logistic and a mean-zero normal with standard deviation $\sqrt{1+\sigma^2}$. Since the distribution of C_{it} is constant across cities, a constant fraction of firms δ_t in each city in any given time period will consider moving. Note that the left-hand side is increasing in σ; more firms move in periods when the variance of the common component of city utilities

Review of Economics and Statistics Just Accepted MS.
https://doi.org/10.1162/rest_a_01381

is higher, and fewer move as the distribution of search costs C_{it} shifts leftward.

The number of firms who are born in city j and *move* to city k in period t is

$$M_{jkt} = N \frac{e^{B_j}}{e^{B_0} + \sum_l e^{B_l}} \delta_t \frac{e^{V_{kt}}}{\sum_l e^{V_{lt}}}$$

That is, the number of firms born in j who move to k in period t is equal to the number of firms born in j times the probability a given firm considers moving in period t times the probability it gets its highest utility draw at that time from city k.[10] We therefore have that

$$\frac{M_{kjt}}{M_{jkt}} = \frac{e^{V_{jt}} e^{B_k}}{e^{V_{kt}} e^{B_j}} = \frac{e^{V_{jt} - B_j}}{e^{V_{kt} - B_k}} = \frac{e^{\bar{V}_{jt}}}{e^{\bar{V}_{kt}}}$$

Letting $W_{jt} = e^{\bar{V}_{jt}}$, we have

$$\frac{M_{kjt}}{M_{jkt}} = \frac{W_{jt}}{W_{kt}}$$

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

$$\sum_j M_{kjt} W_{kt} = \sum_j M_{jkt} W_{jt}$$

and hence

$$\frac{\sum_j M_{jkt} W_{jt}}{\sum_j M_{kjt}} = W_{kt}$$

[10] In periods $t = 2, 3, 4...$, the fact that some firms have already moved once does not affect this formula. It specifically derives the fraction of firms *born* in j who move to k in period t. The decision problem of a given firm on whether to search a second time is, as derived above, independent from whether it has already moved in the past, and the probability k is maximal if it does so is likewise independent of what city the firm currently resides in.

Review of Economics and Statistics Just Accepted MS.
https://doi.org/10.1162/rest_a_01381

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 $\bar{V}_t$. 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 $S_{jkt} = \frac{M_{jkt}}{\sum_n M_{jnt}}$. 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

Review of Economics and Statistics Just Accepted MS.
https://doi.org/10.1162/rest_a_01381

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.[11] Given that the eigenvector represents relative values of W, we can convert into city relative common utilities by $\bar{V}_{jt} = \ln(W_{jt})$, using the definition of W.

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 our assumptions is independent of N, the number of potential entrepreneurs, δ_t, 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 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 $\bar{V}_t$ is based on an eigenvector whose value is constant regardless of the number of non-movers M_{jjt}.[12] 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

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

[12] Since $\bar{V}$ is completely determined by linear equations of the form $\sum_j M_{kjt}W_{kt} = \sum_j M_{jkt}W_{jt}$, the diagonal element M_{jjt} appears as $M_{jjt}W_{jt}$ on both sides and hence cancels out.

Review of Economics and Statistics Just Accepted MS.
https://doi.org/10.1162/rest_a_01381

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 four firms moving in or moving out, and directly check that the matrix of moves is invertible.[13] This constraint binds particularly for LLCs, which have less mobility than corporations. Since cities outside the strongly connected set by definition have very few moves in or out, 98.9% of all interstate corporation moves to the states in our sample nonetheless are to MSAs within this strongly connected set. In all tables, we denote by "N/A" the utility of cities which are dropped because of this restriction.

3.3 Model Implications

The model both permits a ranking of city utilities for firms of differing vintages to be estimated, but also provides an interpretation of the stylized empirical facts of entrepreneurial migration. These stylized facts can be divided into three types: which firms move at all, how the decision of where to move varies by firm age, and how the decision of where to move varies by firm size.

Consider first the decision to move at all. There is a fixed cost of moving C_i which must be overcome to make moving worthwhile, even if most firms were in the absence of that cost mismatched with their highest-utility city. Smaller firms, in terms of lifetime expected profitability, will therefore be less likely to move at any given age. Firms whose expected profitability is more variable across cities, operationalized by σ in the model, are on the other hand more likely to move at any given size or age, since when they plan a move, they get the maximum city utility, not the average. A long theoretical and empirical literature has argued that young firms have more variable growth rates and productivity due to the need

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

Review of Economics and Statistics Just Accepted MS.
https://doi.org/10.1162/rest_a_01381

to learn the best way to run their business (e.g., Jovanovic, 1982). We therefore expect the highest move rates for young firms and those with high expected profits.

Conditional on considering a move, the geographic pattern of migration may vary by firm age. The nursery cities model of Duranton and Puga (2001) combines the insights of Jane Jacobs and Alfred Marshall to argue that young firms benefit from being in an idea-rich, industrially-diverse environment. As firms stabilize their products and business model, they instead are better off being in more specialized cities. If nursery cities help explain entrepreneurial migration, then the idea-rich, diverse cities should have higher utility for young movers than for older ones.[14]

Finally, the idiosyncratic preferences of founders or owners and the direct effect on firm profitability can both drive move decisions. For instance, Guzman (2023) argues that there is a causal benefit to relocating to Silicon Valley for a young firm, and shows that, conditional on firm quality, young founders are more likely to move. If city utility is partly personal to the founder (better weather, greater amenities as in Albouy (2016), lower housing costs, etc.) and partly beneficial to the firm's future profits, firms with lower growth intentions and hence lower expected lifetime profitability will, ceteris paribus, be more likely to move to high-amenity cities rather than nursery cities.

[14]Note that "double movers" may still be rare in our data even if the nursery cities model holds. The reason is that if very few firms move when very young to San Francisco relative to the number born there, for instance, the great majority of San Francisco-based firms who can consider leaving to a more specialized city when old will be ones born in that city. That is, you need to not just be mismatched in both periods, but also to have low enough moving costs in each period.

Review of Economics and Statistics Just Accepted MS.
https://doi.org/10.1162/rest_a_01381

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 in a co-working space, and some time working while traveling. In this case, the location of the firm itself is unclear. Second, 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 commonly-used databases. Finally, 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

Review of Economics and Statistics Just Accepted MS.
https://doi.org/10.1162/rest_a_01381

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 jurisdiction[15] 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. There are two reasons for this definition, one practical and one theoretical. On practical grounds, since data on firm registrations does not include the date the actual principal office was moved, we can only use the date the firm first registered in a state which at a later point shows that state as being the location of the principal office. Theoretically, a company which opens an office in a given state, registering there, but which then performs more hiring and other functions until that state is referred to legally as the principal office, even theoretically should have the initial date of registration as the beginning of the eventual migration. Full details of this process, including how it differs from commercial business registries which generally do not identify startups as young as the

[15]Note 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.

Review of Economics and Statistics Just Accepted MS.
https://doi.org/10.1162/rest_a_01381

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.[16]

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 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).

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. Our "city utility" should therefore be interpreted as the utility to non-regional movers, rather than reflecting, for instance, regional competition for firms. While this could lead to a different migration *rate*

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

Review of Economics and Statistics Just Accepted MS.
https://doi.org/10.1162/rest_a_01381

for larger or smaller states, our empirical approach identifies the relative mover utility of cities using only bilateral moves for each city pair and hence is unaffected by these omissions. 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.[17]

Though business registries appear to be a promising data source for investigating startup

[17]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.

Review of Economics and Statistics Just Accepted MS.
https://doi.org/10.1162/rest_a_01381

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.[18]

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., 2022). The Startup Cartography Project is a project measuring the founding registration of all companies in the United States outside of Delaware, between 1988 to 2014. From this data, we attempt to extract the local and principal address of office for each firm that also has a Delaware jurisdiction. Using this approach, we excluded 15 states in which we did not think we were able to adequately separate the local address from the headquarters. The states we use represent 82% of the US population and 86% of the 50 biggest metropolitan areas by population. See Online Appendix Figure A2 for a visual display of the included states.[19]

[18]For 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.

[19]We also omit the Trenton, NJ and Augusta, ME firms due to idiosyncrasies in how states record firms in these capital cities.

Review of Economics and Statistics Just Accepted MS.
https://doi.org/10.1162/rest_a_01381

4.2 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% have had an IPO and 2.6% have been acquired.[20] Highlighting the growth orientation of these companies, their probability of positive growth outcomes is more than thirty times higher than that of all new firms, as 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, 3.3% of Delaware-registered corporations move to a new state in our data within 2 years, and 5.6% within 5 years. Accounting for moves to states we do not observe as described in Section 2, our estimated overall migration rate for corporations is 6.6%. 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 zero high growth startups, and the median MSA receives one.

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 over our entire 27 year period, 47% of state dyads do not have a single move between them. Panel B describes the much sparser MSA to MSA flows. 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

[20]IPO measures whether the firm joins the NYSE or the NASDAQ as reported by the SDC Global New Issues database. 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.

Review of Economics and Statistics Just Accepted MS.
https://doi.org/10.1162/rest_a_01381

movers conditional on having at least one move between MSA pairs is 3.8. This sparseness highlights the value of our empirical method, which uses network properties rather than just bilateral flows to value cities.

Online Appendix Figure A3 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%.

Figure 3 documents the declining rate of migration over time, including among the highest-quality firms. 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 3 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

Review of Economics and Statistics Just Accepted MS.
https://doi.org/10.1162/rest_a_01381

year sample, for readability we restrict here to MSAs with a population over 1 million.[21] Column 1 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. 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).

Figure 4 shows the relationship between startup births per capita and utility for movers graphically. The top-left panel makes clear that there is no relationship between cities that create many startups per capita and those that are attractive to moving startups. However, this finding hides an intriguing lifecycle pattern. In their first year after being founded, or in the first two years, there is a strong *positive* relationship between cities that birth many

[21]The listing of all cities by utility can be found in Online Appendix Table A2.

Review of Economics and Statistics Just Accepted MS.
https://doi.org/10.1162/rest_a_01381

startups and those that are attractive to movers. Between the third and fifth year after being founded, however, the relationship is precisely the reverse: cities with few births per capita are now the more attractive ones.

Table 4 shows this pattern formally in the first column. The utility of cities to startups moving in their first two years is positively related to the number of startups per capita those cities create, while the opposite pattern holds for later startup moves. To the extent that startups are "mismatched" and must move, they do not move randomly; rather, initially the cities that already had many startups on average benefit from this mobility, whereas as startups become more advanced, they begin to move away from those hubs.

Online 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.

What might explain these empirical regularities? Recall that our theoretical model predicts the following facts. First, younger firms move more since the variation across cities matters more to firms with a less-settled business model. Second, firms with more growth intention move more, since they are more likely to find it worth the cost of switching cities. Third, if the Duranton and Puga (2001) "nursery cities" model holds, young firms optimally locate in places with a diverse set of ideas and industries, while they move to lower-cost, more-specialized cities as their business develops. Fourth, if founders consider both pecuniary and non-pecuniary factors, aspects of cities that affect pure economic return should matter more to founders with stronger growth intention. And of course, only firms so "mismatched" with their original city move at all given the cost of doing so.

Online Appendix Table A5 shows that the first two hypotheses hold. The fraction of startups that move is highest the year the firms are founded, and monotonically falls thereafter. However, corporations of any age are much more likely to move than LLCs, those

Review of Economics and Statistics Just Accepted MS.
https://doi.org/10.1162/rest_a_01381

holding IP at birth (in addition to many other measures of growth intention at birth) are as well, and later movers are also much more likely to be acquired or IPO than Delaware-registered firms who either never move or who move when very young. The differences are substantial: of Delaware-registered firms who do not change states in their first five years, 43% are corporations rather than LLCs, while among those who move in their first year, 57% are corporations, and among those moving in year five, 65%.

Columns 2, 3 and 4 in Table 4 test the nursery cities hypothesis, in two ways. First, we measure idea diversity using the four-digit employment HHI of each MSA, where a higher HHI means the city has employment concentrated among fewer sectors. While industrial concentration is strongly negatively associated with the number of startups born per capita, there is no relationship between HHI and utility for startups who move in the first two years after founding, and a positive relationship between industrial concentration and utility for startups who move later. Likewise, while patenting per capita is strongly positively associated with startup births, it is negatively associated with utility for late startups movers. For both measures of "idea diversity", the nursery cities pattern holds.

Columns 5 through 8 in Table 4 examine financial motives for moving by regressing total state-level tax rates as computed by Moretti and Wilson (2017) against startup births and mover utility. While high tail taxes (in this case, 95th percentile income taxes) are not associated with either less entrepreneurship or lower utility for early movers, later movers show a large, negative reaction to these tail taxes. Figure 5 shows this relationship graphically. Online Appendix Table A7 shows that including corporate tax rates makes the negative reaction of late movers to high taxes even more stark: high tail individual tax rates and high corporate tax rates independently repel late-moving startups.

While Figure 6 shows that LLCs are also deterred from moving to cities with high tail taxes, Online Appendix Table A8 shows that the relationship between pecuniary factors and Delaware-registered LLCs is dulled compared to that of corporations. The nursery cities relationships are not statistically significant and not evident even in the point estimates,

Review of Economics and Statistics Just Accepted MS.
https://doi.org/10.1162/rest_a_01381

and when it comes to taxes, LLCs if anything respond more strongly to median tax rates than those at the right tail. Appendix Table A10 regresses city utility on purely non-pecuniary factors such as sunshine and the Albouy (2016) "quality of life" index derived from individual rather than corporation moves. While these factors show no relationship with mover utility for corporations, utility for moving LLCs is strongly associated with sunshine, warmer weather, and higher quality of life. Combined with the tax evidence, this is consistent with the idea that founders with less growth intention, who initially form their ventures as LLCs, react less to financial factors and more to other factors relevant to the personal utility of their founders.

How important are startup moves to the overall number of startups in a city? For cities that generate many startups, net movement is relatively unimportant: San Jose loses a net 25 high-growth startups during a period in which they create almost 9,000. However, startup births are highly skewed, hence startup mobility can be quite important. The median city in startup births per capita in our data, San Antonio, would move ahead of ten more cities in total post-move startups per capita if they had the per capita attractiveness to movers of Austin, and behind seven cities if they had that of New Orleans. Put another way, even though the vast majority of startups don't move and the big startup hubs are driven much more by firm creation than firm mobility, a San Antonio that could attract startups as well as Austin would see roughly 20 percent more age-5 startups than a San Antonio which attracted startups at the rate of New Orleans. That is, while cities like Boston and Mountain View may barely notice that companies like Facebook and Tableau left, a city like Albuquerque or Houston would absolutely notice if they arrived.

Before concluding, let us consider three threats to our empirical approach: that moves are idiosyncratic, or that overall net movement is driven by vertical or horizontal sorting.

Consider first idiosyncratic moves. Of course, some startup moves are heavily influenced by the idiosyncratic preferences of founders; for example, Microsoft's relocation to Seattle appears to be partially influenced by the fact that Bill Gates and Paul Allen wanted to

Review of Economics and Statistics Just Accepted MS.
https://doi.org/10.1162/rest_a_01381

be close to their families. 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 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 $H_0 : E[\frac{Moves\ In_i}{Moves\ Out_i}] = 1$ for any given city. Empirically, this set of hypotheses can be tested with a joint Fisher Exact Chi-squared Test. However, it is straightforward to see that even individual cities have combinations of moves in and moves out that are wildly unlikely to be the result of idiosyncratic movement alone. For example, Dallas has 453 moves in and 215 moves out, violating H_0 at p<.00001.

Second, consider the relationship between city utility and an estimate of the vertical quality at birth of firms moving to or from each city. For example, while places with high startup costs such as the Bay Area may on net lose startups, 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 investigate vertical sorting, we replicate 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.[22] Online Appendix Table A6 shows that firms which move to startup hubs, including those that move to startups hubs in their third to fifth year after being founded, are higher quality than

[22]For all non-movers born before 2012, we run a logit model of a binary measure of equity events on observables for whether a firm is born a corporation, has a short name, is eponymous, has a patent, has a trademark, or is estimated to be part of certain industries based on firm name. Predictions from this model havean out of sample ROC score or 0.80. Estimated quality is then the predicted out of sample probability of this model.

Review of Economics and Statistics Just Accepted MS.
https://doi.org/10.1162/rest_a_01381

those who move to non-hubs. That said, Online Appendix Table A12 regresses the quality of startups that arrive on the quality of those that leave controlling for the average quality of all firms born in that city and finds, controlling for level of growth startups per capita in a MSA, no relationship between the quality of leavers versus stayers. That is, startup hubs like Silicon Valley do in fact both create and attract high quality startups, but the startups they lose are also disproportionately high quality. In short, we do not find evidence that our primary results are driven by quality-based vertical sorting.

Finally, Appendix Figure A1 shows that our city utilities overall are highly correlated with city utilities estimated using only companies in the health sector, IT sector, services sector, or high tech sector. As in Guzman and Stern (2020), we predict industry from a firm's name. In a model of pure horizontal sorting, the rank of *within-industry* utilities should vary. Instead, cross-industry factors cause a strong correlation between the overall city rank and the rank within industry. Of course, we are only able to identify industries in broad classes, and with error, since we do not directly observe industry.

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 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

Review of Economics and Statistics Just Accepted MS.
https://doi.org/10.1162/rest_a_01381

Francisco, and Boston produce a wildly disproportionate number of high-growth startups per capita, as do university towns, startup moves are not a case of "the rich get richer". While very young startups are more likely to move to cities that birthed many startups, those between two and five years after founding are much more likely to move to low-tax, not-terribly-Bohemian Sunbelt cities. This pattern is most evident for the most growth-oriented firms: those registered as Delaware corporations at birth are particularly likely to shift their headquarters to boring, business-friendly locations, while LLCs, perhaps accounting for non-pecuniary tastes of their founders, move to sunny, high-amenity destinations. And while this pattern holds over our full sample period, the "attractive cities" to high-quality startups are not set in stone. As can be seen in Online Appendix Table A4, Las Vegas, Nashville, Austin and San Antonio have become relatively more attractive post-2001 compared to the 1990s, while Minneapolis, Richmond, Houston and Denver have become less so.

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 identify 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.

These results 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 previously believed. 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.

Review of Economics and Statistics Just Accepted MS.
https://doi.org/10.1162/rest_a_01381

References

John M Abowd, Francis Kramarz, and David N Margolis. High wage workers and high wage firms. *Econometrica*, 67(2):251–333, 1999.

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. Reprint of the startup cartography project: measuring and mapping entrepreneurial ecosystems. *Research Policy*, 51(9):104581, 2022.

Reiner Braun and Stefan Weik. Foreign venture capital and the exodus of start-ups: Evidence from headquarters relocations. *Working Paper*, 2021.

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.

Annamaria Conti and Jorge A Guzman. What is the us comparative advantage in entrepreneurship? evidence from israeli migrantion to the united states. *Review of Economics and Statistics*, 105(3):528–44, 2023.

Michael S Dahl and Olav Sorenson. The embedded entrepreneur. *European management review*, 6(3):172–181, 2009.

Review of Economics and Statistics Just Accepted MS.
https://doi.org/10.1162/rest_a_01381

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.

William H Frey. U.s. migration still at historically low levels, census shows, 2017.

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.

Jorge Guzman. Go west young firm: The impact of startup migration on the performance of migrants. *Management Science*, forthcoming, 2023.

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`.

Review of Economics and Statistics Just Accepted MS.
https://doi.org/10.1162/rest_a_01381

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.

Boyan Jovanovic. Selection and the evolution of industry. *Econometrica*, 50(3):649–670, 1982. ISSN 00129682, 14680262. URL `http://www.jstor.org/stable/1912606`.

William R. Kerr and Frederic Robert-Nicoud. Tech clusters. *Journal of Economic Perspectives*, 34(3):50–76, August 2020.

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.

AnnaLee Saxenian. *Regional networks: industrial adaptation in Silicon Valley and route 128*. Citeseer, 1994.

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.

Isaac Sorkin. Ranking firms using revealed preference. *The quarterly journal of economics*, 133(3):1331–1393, 2018.

Review of Economics and Statistics Just Accepted MS.
https://doi.org/10.1162/rest_a_01381

Figure 1: New Startups and Net Startup Moves by MSA

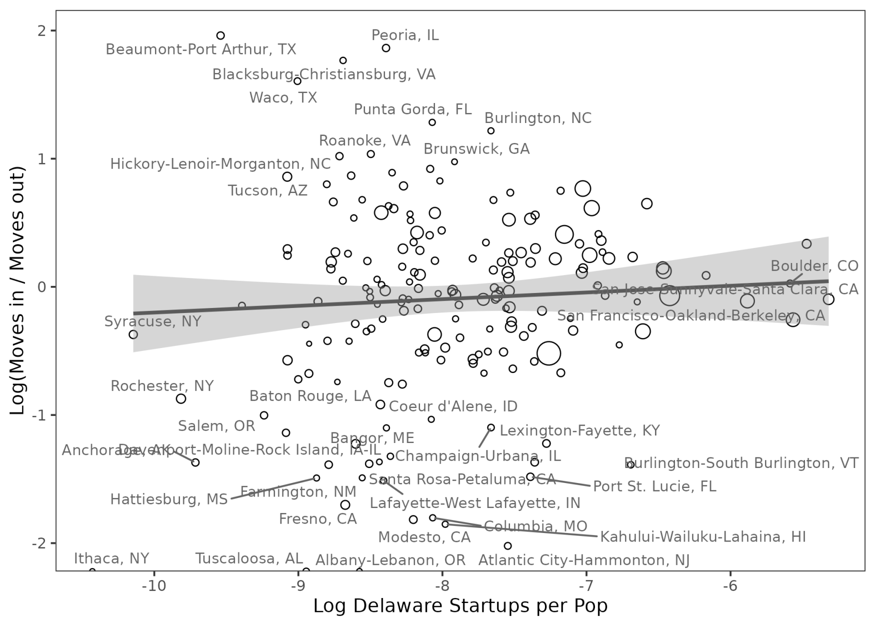
A. New Firm Formation by CBSA

B. Net Startup Migration by CBSA

The top map plots the number of startups per capita by MSA, where darker shading signifies more startups, and bubble size signifies MSA population. As in the remainder of the results, "startup" refers to a corporation or LLC registered in Delaware at birth. The bottom map plots startup moves, where the size of the circle is the number of total movers that move into a metropolitan area, and the color of the circle represents the ratio of the moves in over the moves out. Darker cities have a higher number of moves in than moves out, while lighter cities are the opposite.

Review of Economics and Statistics Just Accepted MS.
https://doi.org/10.1162/rest_a_01381

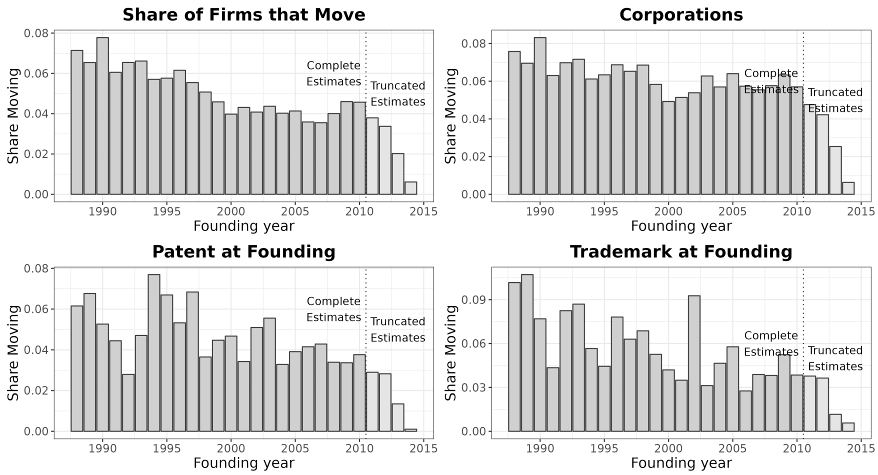
Figure 2: Estimated Migrant Value of Cities vs Local Ecosystem Strength



The ratio of high-quality startups moving in versus moving out of an MSA is uncorrelated with the number of high-quality startups born in that city. Bubble size represents city population.

Review of Economics and Statistics Just Accepted MS.
https://doi.org/10.1162/rest_a_01381

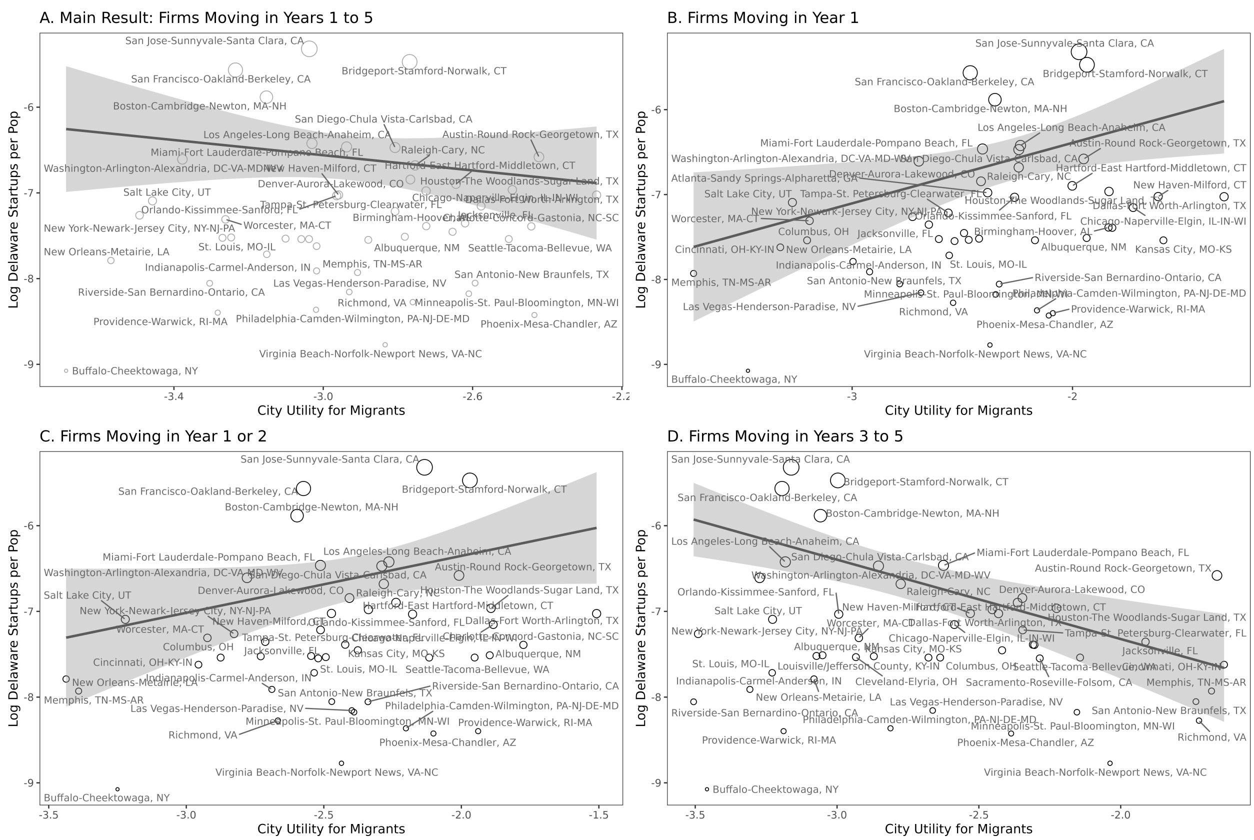
Figure 3: Migration Rate by Founding Year, 1988-2014



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.

Review of Economics and Statistics Just Accepted MS.
https://doi.org/10.1162/rest_a_01381

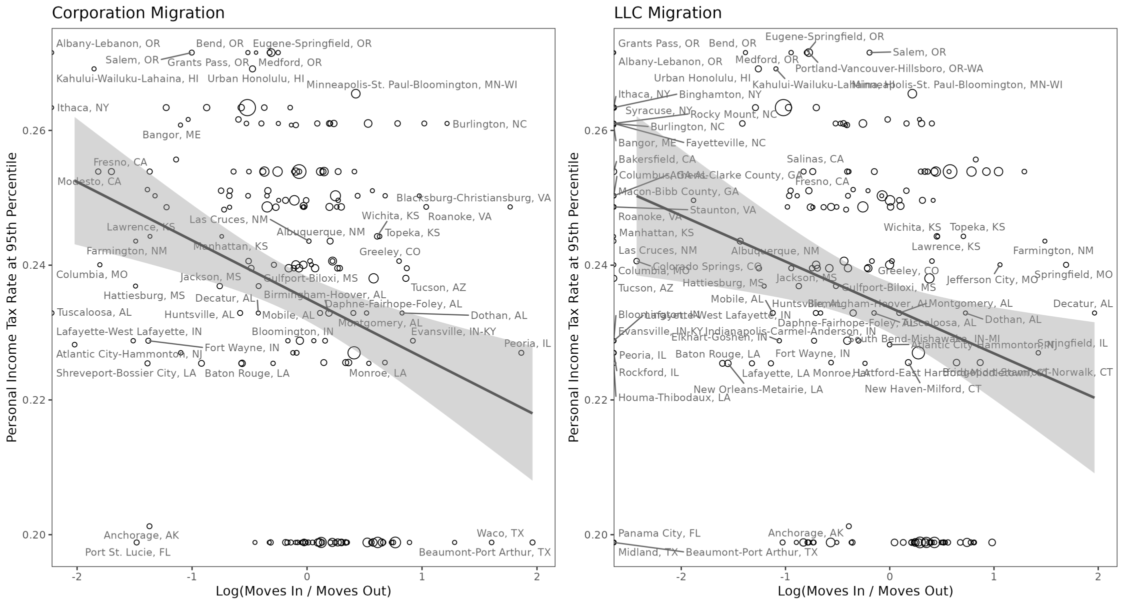
Figure 4: Migrant City Utility Across Migration Age



The figure plots the estimated relative city utility for moving corporations based on the age at which they move. Panel A is all movers aged 0-5 years, Panels B through D split these into smaller year ranges. The fitted line is weighted by the ecosystem startup intensity (startups per capita). Bubble indicates startups founded in each city per capita.

Review of Economics and Statistics Just Accepted MS.
https://doi.org/10.1162/rest_a_01381

Figure 5: Net Migration Rates and Taxes



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 at the 95th percentile of income in that state, estimated by Moretti and Wilson (2017). We observe a large negative correlation between both variables.

Review of Economics and Statistics Just Accepted MS.
https://doi.org/10.1162/rest_a_01381

Table 1: Summary Statistics

Panel A: Corporations (N=181663)

Statistic	Mean	St. Dev.
Incorporation Year	2,001.819	7.470
IPO	0.005	0.071
Acquired	0.026	0.159
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

Panel B: LLCs (N=237307)

Statistic	Mean	St. Dev.
Incorporation Year	2,006.634	5.320
IPO	0.0001	0.010
Acquired	0.004	0.065
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

Panel C: Estimated 5-year U.S. Migration Rates

Corporations	0.066
LLCs	0.032

Review of Economics and Statistics Just Accepted MS.
https://doi.org/10.1162/rest_a_01381

Table 2: Summary Statistics for Migrant Flows Data

Panel A: State to State Migration Flows

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

Panel B: CBSA to CBSA Migration Flows

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

Review of Economics and Statistics Just Accepted MS.
https://doi.org/10.1162/rest_a_01381

Table 3: Estimated Utility for Large US Cities (Population over 1 million in 2010)

Log Utility	Rank (age: 1-5)	CBSA Name	Moves In	Moves Out	Moves In LLC	Moves Out LLC	LLC Rank (age: 1-5)
-2.2455	1	Dallas-Fort Worth-Arlington, TX	453	215	324	241	16
-2.3842	2	Phoenix-Mesa-Chandler, AZ	94	53	47	32	9
-2.4	3	Austin-Round Rock-Georgetown, TX	166	88	96	74	8
-2.436	4	Charlotte-Concord-Gastonia, NC-SC	108	63	65	83	29
-2.4716	5	Houston-The Woodlands-Sugar Land, TX	376	205	282	184	17
-2.5033	6	Seattle-Tacoma-Bellevue, WA	145	86	43	76	32
-2.5524	7	Chicago-Naperville-Elgin, IL-IN-WI	471	311	341	261	11
-2.5824	8	San Antonio-New Braunfels, TX	45	26	25	20	25
-2.5872	9	Minneapolis-St. Paul-Bloomington, MN-WI	90	59	82	65	18
-2.6072	10	Jacksonville, FL	48	35	49	19	1
-2.6117	11	Nashville-Davidson–Murfreesboro–Franklin, TN	96	72	86	61	10
-2.6395	12	Hartford-East Hartford-Middletown, CT	89	63	29	15	6
-2.708	13	Atlanta-Sandy Springs-Alpharetta, GA	363	286	217	239	26
-2.7352	14	Richmond, VA	26	19	22	23	23
-2.739	15	Raleigh-Cary, NC	98	77	33	24	14
-2.7569	16	Denver-Aurora-Lakewood, CO	237	191	114	177	30
-2.77	17	Birmingham-Hoover, AL	51	42	36	51	33
-2.7925	18	Tampa-St. Petersburg-Clearwater, FL	95	78	88	42	4
-2.7928	19	San Diego-Chula Vista-Carlsbad, CA	153	134	217	77	2
-2.8196	20	Virginia Beach-Norfolk-Newport News, VA-NC	19	16	9	16	35
-2.8666	21	Memphis, TN-MS-AR	35	35	43	41	22
-2.8958	22	Sacramento-Roseville-Folsom, CA	32	28	38	15	3
-2.898	23	Las Vegas-Henderson-Paradise, NV	62	56	33	78	39
-2.9264	24	Miami-Fort Lauderdale-Pompano Beach, FL	349	314	339	239	12
-2.9466	25	Orlando-Kissimmee-Sanford, FL	75	68	73	58	20
-2.9866	26	Cincinnati, OH-KY-IN	52	55	24	43	36
-2.9909	27	Columbus, OH	52	49	27	42	38
-3.0023	28	Indianapolis-Carmel-Anderson, IN	47	49	31	47	27
-3.0143	29	Los Angeles-Long Beach-Anaheim, CA	507	544	663	374	7
-3.0493	30	San Jose-Sunnyvale-Santa Clara, CA	213	238	53	35	15
-3.0505	31	Kansas City, MO-KS	74	78	18	36	43
-3.0929	32	Cleveland-Elyria, OH	45	52	39	50	31
-3.1366	33	Boston-Cambridge-Newton, MA-NH	487	548	253	253	24
-3.1437	34	St. Louis, MO-IL	80	90	8	9	19
-3.2156	35	San Francisco-Oakland-Berkeley, CA	336	433	242	154	13
-3.2365	36	Louisville/Jefferson County, KY-IN	43	58	21	45	41
-3.2478	37	Portland-Vancouver-Hillsboro, OR-WA	100	134	46	100	37
-3.2753	38	Providence-Warwick, RI-MA	12	14	23	21	21
-3.2799	39	Riverside-San Bernardino-Ontario, CA	27	38	23	10	5
-3.3593	40	Washington-Arlington-Alexandria, DC-VA-MD-WV	257	362	140	181	28
-3.4623	41	New York-Newark-Jersey City, NY-NJ-PA	615	1038	263	731	42
-3.4623	42	Salt Lake City, UT	52	72	22	47	40
-3.5586	43	New Orleans-Metairie, LA	35	63	14	66	45
-3.6558	44	Buffalo-Cheektowaga, NY	11	19	1	4	44
-3.8411	45	Rochester, NY	5	11	2	4	34

Review of Economics and Statistics Just Accepted MS.
https://doi.org/10.1162/rest_a_01381

Table 4: What Predicts City Utility?

	Baseline	**Nursery Cities**			**Income Taxes**			
	Migrant City Utility	City Entrepreneurship	*Migrant* City Utility	*Migrant* City Utility	City Entrepreneurship	City Entrepreneurship	*Migrant* City Utility	*Migrant* City Utility
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Growth Startups per Capita	0.183**							
	(0.066)							
Growth Startups per Capita × Later Movers (Years 3-5)	−0.315***							
	(0.090)							
Industry Concentration (HHI)		−0.108**	−0.051					
		(0.051)	(0.036)					
Industry Concentration (HHI) × Later Movers (Years 3-5)			0.095*					
			(0.051)					
Patenting per Capita		0.503***		0.049				
		(0.067)		(0.052)				
Patenting per Capita × Later Movers (Years 3-5)				−0.171*				
				(0.088)				
Personal Income Tax (95th)					4.145		−2.979	
					(3.591)		(2.512)	
Personal Income Tax (95th) × Later Movers (Years 3-5)							−8.751**	
							(3.717)	
Personal Income Tax (50th)						−11.212*		−6.412*
						(5.946)		(3.469)
Personal Income Tax (50th) × Later Movers (Years 3-5)								−1.521
								(5.546)
Observations	138	136	136	138	138	138	138	138
R^2	0.198	0.401	0.150	0.151	0.011	0.038	0.271	0.167

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 parentheses. All regressions are OLS. Significance denoted as *p<0.1; **p<0.05; ***p<0.01