NBER WORKING PAPER SERIES

IMMIGRATION AND ENTREPRENEURSHIP IN THE UNITED STATES

Pierre Azoulay
Benjamin Jones
J. Daniel Kim
Javier Miranda

Immigration and Entrepreneurship in the United States
Pierre Azoulay, Benjamin Jones, J. Daniel Kim, and Javier Miranda
NBER Working Paper No. 27778
September 2020
JEL No. J15,L26,M13,O3

ABSTRACT

Immigration can expand labor supply and create greater competition for native-born workers. But immigrants may also start new firms, expanding labor demand. This paper uses U.S. administrative data and other data resources to study the role of immigrants in entrepreneurship. We ask how often immigrants start companies, how many jobs these firms create, and how these firms compare with those founded by U.S.-born individuals. A simple model provides a measurement framework for addressing the dual roles of immigrants as founders and workers. The findings suggest that immigrants act more as "job creators" than "job takers" and that non-U.S. born founders play outsized roles in U.S. high-growth entrepreneurship.

Pierre Azoulay
MIT Sloan School of Management
100 Main Street, E62-487
Cambridge, MA 02142
and NBER
pazoulay@mit.edu

Benjamin Jones
Northwestern University
Kellogg School of Management
Department of Management and Strategy
2211 Campus Drive
Evanston, IL 60208
and NBER
bjones@kellogg.northwestern.edu

J. Daniel Kim
University of Pennsylvania
Wharton School
3620 Locust Walk
2029 Steinberg-Dietrich Hall
Philadelphia, PA 19104
jdkim@wharton.upenn.edu

Javier Miranda
U.S. Bureau of the Census
Economy-Wide Statistics Division
4600 Silver Hill Road
Washington, DC 20233
javier.miranda@census.gov

1 Introduction

Economic perspectives of immigration often emphasize its role in expanding labor supply (e.g., Isaac 1947, Borjas 1994, Dustmann et al. 2016). From this starting point, immigrants may primarily appear to compete with native workers, leading to reduced employment or lower wages (e.g., Altonji and Card 1991, Borjas 2003). Native workers and their representatives may then oppose immigration on the grounds that immigrants "take jobs" (e.g., Reder 1963, Briggs 2001). However, this perspective, while common in economic research and powerful in policy, does not appear to be the whole story. For example, studies of immigration shocks—including the Mariel Boatlift and others—often do not find negative effects on local wages (Card 1990, Hunt 1992, Friedberg 2001). And studies of mass migration to the United States have found substantial and persistent gains in per-capita income in regions that experience greater immigration (Tabellini 2020; Sequeira et al. 2020). The tension between a labor supply orientation and the broader empirical findings suggests that additional economic forces are at work.

This paper works to fill in the picture through the lens of entrepreneurship. We consider how immigrants not only expand labor supply (as workers) but also expand labor demand (as founders of firms). Using administrative data for the U.S. economy, we study the extent to which immigrants start new firms, and we study the employment in the firms they found. By looking in a more comprehensive manner at the U.S. economy, the analysis helps balance the ledger in assessing immigrants' economic roles.

To provide a coherent analysis of supply and demand, we first present a simple general equilibrium model of the economy. This model, building on Lucas (1978), considers how individuals sort into workers and founders. We extend the model to consider the role of immigration. Intuitively, if immigrants start many firms but these are small firms, their effect on labor demand will still be small, so that immigration mainly expands labor supply and on net reduces wages. However, if immigrant-founded firms tend to be large firms, then immigration may substantially expand labor demand, raising wages. The theory makes this explicit. The model also generates specific predictions for the distribution of firm size, including Zipf's Law, that can be brought to the data.

The empirical work then builds from three data sets. First, we use administrative records to study every firm founded in the U.S. between 2005 and 2010. Second, we use the U.S. Census's Survey of Business Owners to study a representative sample of all U.S. firms. Third, we examine the Fortune 500, allowing us to focus on the very largest firms in the U.S. economy. In all cases, we code founders as either U.S. born or immigrant based on their place of birth. Using any of these data sets, we find similar results. First, reflecting existing research, immigrants start firms at higher rates than native-born individuals do (e.g., Kerr and Kerr 2020). Second, immigrants do not simply start small firms. Rather, they tend to start more firms at every size, compared to U.S.-born individuals. This is the key finding of the paper. There is effectively a "right shift" in entrepreneurship for immigrants, with immigrants playing relatively large roles as employers rather than employees, compared to U.S. born individuals.

Overall, the findings suggest that immigrants appear to "create jobs" (expand labor demand) more than they "take jobs" (expand labor supply) in the U.S. economy. By studying immigrants as both entrepreneurs and workers, one produces a fuller picture of the impact of immigration. The new facts can help resolve existing puzzles where empirical evidence, including natural experiments and longer-run historical analysis, often suggests more positive economic effects from immigration than labor-supply oriented perspectives produce.

The paper proceeds as follows. Section 2 further motivates our study of immigrant entrepreneurship. Section 3 presents a simple model, generating key intuitions and clarifying basic empirical constructs. Section 4 presents the data and empirical measures. It further presents our key findings. Section 5 considers interpretations and extensions, including patenting behavior. Section 6 concludes.

2   The Immigrant Entrepreneur

While researchers often study immigrants as workers, a recent stream of scholarship emphasizes the role of immigrants as innovators and entrepreneurs. For example, immigrants currently represent about 14% of the U.S. workforce but account for closer to one-quarter of U.S. patents (e.g., Bernstein et al. 2019). Related, immigrants to the U.S. are also disproportionately likely to hold STEM degrees (Kerr and Kerr 2020).

Immigrants further appear to be highly entrepreneurial. Immigrants start firms at higher rates than native-born individuals in several countries (Fairlie and Lofstrom 2015), with recent evidence suggesting that approximately 25% of recent start-ups in the U.S. are founded by immigrants (Kerr and Kerr 2020). At the same time, such entrepreneurial tendencies may largely reflect businesses with limited growth prospects, perhaps because immigrants are "pushed" into entrepreneurial activity due to poor labor market opportunities (Light and Roach 1996, Constant and Zimmerman 2006). On the other hand, evidence also points to the substantial presence of immigrant founders in Silicon Valley (e.g., Saxenian 2002), and examples such as Alexander Graham Bell and Sergei Brin suggest that immigrants have started firms that grow to employ large numbers of people, with large effects on the economy.

Ultimately, the economic effects of immigration will depend not just on whether immigrants start firms, but on how successful these firms tend to be. In this paper, we will provide systematic evidence on this question. Intuitively, examining the size distribution of businesses, and comparing this for immigrant and native founders, can shed light on the impact of immigrant entrepreneurship. We will examine the size distributions using administrative data and other data resources. To provide intuition and make the measures precise, we first develop a simple conceptual framework.

3   Model

Immigrants can expand both labor supply (as workers) and labor demand (as employers). Immigrants, as earners of income, also create more demand for final goods. The following model builds on Lucas (1978) to introduce such general equilibrium reasoning. We extend this classic model to consider the net effect of immigration on the economy and provide explicit constructs that can be examined empirically.

3.1   Assumptions

Let there be N people in the economy, where individuals can choose to either work in a firm ("workers") or start a firm ("founders"). Each person is endowed with 1 unit of labor. Each person is also endowed with some level of entrepreneurial acumen. The entrepreneurial acumen for

individual i is $a_i \geq 0$, which is distributed $f(a)$. The (endogenous) number of founders is E and workers is L, where $E + L = N$

Firms maximize profits. They produce with a decreasing returns-to-scale technology, and productivity depends on the entrepreneur's skill. These features allow for positive profits and a size distribution of firms in equilibrium. Specifically, a firm's output is

$$y_i = a_i l_i^\beta \tag{1}$$

where $\beta \in (0,1)$ and l_i is the labor employed. The profit maximization problem is

$$\pi_i^* = \arg\max_{l_i} [y_i - w l_i] \tag{2}$$

where the final good price is taken as numeraire (there is only one type of output).

Individuals choose their career to maximize income. The individual's choice is to work for a wage, w, or start a firm and earn profit π_i^*. Individuals choose to become entrepreneurs if $\pi_i^* \geq w$ and choose to be workers otherwise. Utility is strictly increasing in consumption of the final good. Individual-level consumption is thus equated to individual-level income, and total consumption is equated to GDP.

Finally, we consider two sub-populations, indexed $j \in \{0, 1\}$, to represent the native-born and immigrants, respectively. The total population is partitioned as $N = N_0 + N_1$ and we similarly partition $L = L_0 + L_1$ and $E = E_0 + E_1$. The distribution of entrepreneurial acumen for each group is $f_0(a)$ and $f_1(a)$. The overall distribution of acumen in the economy is the summation of these two sub-population distributions, each weighted by its population share. Results below will (eventually) specialize to consider Pareto distributions,

$$f_j(a) = \frac{\gamma a_j^\gamma}{a^{\gamma+1}}, \ \gamma > 0, \ a \geq a_j \tag{3}$$

where the parameter a_j acts as a distributional shifter.

3.2 Equilibrium Results

We emphasize two sets of results, helping frame the empirical work to follow. The first set provides more general statements about equilibrium outcomes in light of immigration. The second set considers Pareto distributions of entrepreneurial talent, which provide a close match to the data.

4

To solve for the equilibrium allocation, we have firm-level profit maximization and the individual career decision. These are the choices in the economy. We close the model through the resource constraints, which are the total available population and, most importantly, the distributions of entrepreneurial acumen for the native-born and for immigrants.

First, from profit maximization, profits are strictly increasing in the individual acumen, a_i.[1] The entrepreneurial choice decision then implies a unique threshold value a^*, where individuals choose entrepreneurship if $a_i \geq a^*$ and choose to be workers otherwise. This threshold level of acumen is

$$a^* = \frac{w}{\beta} \left(\frac{\beta}{1-\beta} \right)^{1-\beta} \tag{4}$$

for any distribution of talent, $f(a)$. This equilibrium condition provides a monotonically increasing relationship $a^*(w)$. This produces an upward sloping labor supply relationship; namely, a higher wage means that more people choose to be workers.[2]

Second, the share of entrepreneurs in the population is

$$\frac{E^*}{N} = \int_{a^*}^{\infty} f(a_i) \, da_i \tag{5}$$

which is decreasing in a^*. Fewer founders means less labor demand, other things equal, leading to a second and decreasing relationship between a^* and w. The above two conditions together can thus pin down the equilibrium wage and entry threshold decisions.

Aggregates follow by adding up the firm level variables. Total output per capita is

$$\frac{Y^*}{N} = \int_{a^*}^{\infty} y_i^* f(a_i) \, da_i \tag{6}$$

where the firm-level outputs follow from profit maximization. We can similarly add up firm-level labor demands, l_i^*, and profits, π_i^* to produce aggregate labor, L^*, and aggregate profit, Π^*.[3] Aggregate GDP is constant returns to scale in the population, conditional on a fixed distribution of talent. Equating total income to GDP, it follows, for any $f(a)$, that the labor share of total income is $wL/Y = \beta$ and the aggregate profit share of total income is $\Pi/Y = 1 - \beta$.

---

[1]Firm-level outcomes follow from profit-maximization, (2), given the wage. These are the labor demands $l_i^* = \left(\frac{\beta a_i}{w} \right)^{\frac{1}{1-\beta}}$, outputs $y_i^* = a_i^{\frac{1}{1-\beta}} \left(\frac{\beta}{w} \right)^{\frac{\beta}{1-\beta}}$, and profits $\pi_i^* = a_i^{\frac{1}{1-\beta}} (1-\beta) \left(\frac{\beta}{w} \right)^{\frac{\beta}{1-\beta}}$.

[2]Rather than a choice between work and leisure, here we mean the choice between being a worker (L) and a business owner (E). The phrases "labor supply" and "labor demand" in this construct refer to workers.

[3]The firm-level outcomes all aggregate similarly in the term $a_i^{\frac{1}{1-\beta}}$ and their aggregates are linear in population size.

5

3.2.1 Immigration and Equilibrium

The influence of immigration can by understood by analyzing how immigration shifts the overall distribution $f(a)$. Consider three informative cases.

1. Immigration causes no shift in the distribution $f(a_i)$.

2. Immigration creates a left shift in the distribution $f(a_i)$ such that $f(a_i)$ decreases for all $a_i \geq a^*$.

3. Immigration creates a right shift in the distribution $f(a_i)$ such that $f(a_i)$ increases for all $a_i \geq a^*$.

We can then produce the following results for income in the economy.

Proposition 1. *The threshold for entrepreneurial entry, a^*, is unchanged in case 1, decreasing in case 2, and increasing in case 3. The equilibrium wage (w^*), GDP per capita (Y^*/N), and total profits per capita (Π^*/N) move in the same direction as a^*.*

Proof. See appendix. □

These results are intuitive. Take the first case. Here immigrants are just like the native born. They split into workers and entrepreneurs just like the native born do, and immigrants have no net effect on the equilibrium between labor supply and demand. Immigrants increase the scale of the economy, by increasing total population, but they don't change wages or other outcomes per person.

In the second case, immigrants tend to have less entrepreneurial acumen. Compared to the native born, a relatively high share of immigrants will become workers (pushing out labor supply) rather than entrepreneurs (pushing out labor demand), and equilibrium wages will fall. This case corresponds to often-expressed fears that immigration will worsen wages for native workers.

Finally, consider the third case. Here immigrants tend to have more entrepreneurial acumen than the native born. Although many immigrants may become workers, a relatively high share will

now become business founders, so that the labor demand effect outweighs the labor supply effect, causing the wages of native workers to rise. More generally, GDP per capita and profits per capita will also rise.[4]

An equilibrium with explicit solutions follows using Pareto distributions. To encapsulate these results, first define a constant

$$a_m = [a_0^\gamma (N_0/N) + a_1^\gamma (N_1/N)]^{1/\gamma}.$$

This constant is a weighted combination of the Pareto floor parameters (see (3)) for the native-born and immigrant distributions of entrepreneurial acumen. Further, consider the (natural) case where both the native-born and immigrant populations produce some workers and some entrepreneurs.[5]

Proposition 2. *With the Pareto distributions of entrepreneurial acumen (3), the economy-wide equilibrium outcomes are*

$$a^* = a_m \left[\frac{\gamma - 1}{\gamma (1 - \beta) - 1} \right]^{\frac{1}{\gamma}} \tag{7}$$

$$w^* = a_m \left[\frac{\gamma - 1}{\gamma (1 - \beta) - 1} \right]^{\frac{1}{\gamma}} (1 - \beta)^{1-\beta} \beta^\beta \tag{8}$$

$$E^* = \left(1 - \frac{\gamma \beta}{\gamma - 1} \right) N \tag{9}$$

$$Y^* = a_m \left[\frac{\gamma - 1}{\gamma (1 - \beta) - 1} \right]^{\frac{1}{\gamma}} \beta^\beta (1 - \beta)^{1-\beta} \frac{\gamma}{\gamma - 1} N \tag{10}$$

and other economy-wide entities follow as $\Pi^* = (1 - \beta) Y^*$ and $L^* = N - E^*$.

Proof. See appendix. □

In the empirical work, we will separately examine the entrepreneurial outcomes for native born and immigrant workers. The first result concerns the rate of entrepreneurship.

---

[4]This third case is not, however, a Pareto improvement without a transfer. Although workers are better off, and although total profit per capita goes up, an individual native-born entrepreneur sees his/her profit fall. This follows because, given that entrepreneur's a_i, the profits of his/her firm are declining when the wage increases. That said, from an inequality point of view, this is a case where the individuals who see less income (without a corrective transfer) are those who are relatively well off.

[5]That is, we consider the case where the equilibrium a^* is finite and strictly greater than the floor parameters for both acumen distributions, so that workers and founders emerge within each group. This is of course the natural empirical case.

Corollary 1. *The equilibrium rate of entrepreneurship within a given group is*

$$\frac{E_j}{N_j} = \left(\frac{a_j}{a_m}\right)^\gamma \left(1 - \frac{\gamma\beta}{\gamma - 1}\right) \tag{11}$$

Proof. See appendix. $\qquad\square$

The second result concerns the distribution of firm size. Again, we look separately within the native born and within immigrants. Specifically, we count the number of firms of a given size from the group's founders, normalized by the size of that group's population. We can consider firm size either in terms of employment or output, and here we use employment. Define the count of firms of a given employment size as $n_j\left(l_i^*\right)$ within group j.

Corollary 2. *The (log) fraction of people of type j who start a firm of size l_i^* is*

$$\log \frac{n_j\left(l_i^*\right)}{N_j} = \log C - \left(\gamma\left(1 - \beta\right) + 1\right) \log l_i^* \tag{12}$$

where $C = \left(\frac{a_j}{a_m}\right)^\gamma \left(1 - \frac{\gamma\beta}{\gamma-1}\right)\left(1 - \beta\right)\gamma\left(\frac{1-\beta}{\beta}\right)^{-\gamma(1-\beta)}$ and the smallest firm has size $l_i^* = \frac{\beta}{1-\beta}$.

Proof. See appendix. $\qquad\square$

The firm size distribution is thus a power law—Zipf's Law (Axtell 2001). Namely, the slope of the log frequency on the log firm size is a negative constant. The Pareto distribution approach may thus be useful in matching Zipf's Law in the data. Moreover, through an entrepreneurial lens, differences in these normalized frequency distributions can inform the role of immigration. Immigration produces a right shift in the economy's entrepreneurial acumen distribution if $a_1 > a_0$. This feature will be seen as a right shift in the firm size frequency per group member, (12), and an increased rate of entrepreneurship, (11), when comparing the immigrant and native populations. From Proposition 1, such a right shift increases the wages of domestic workers, as well as GDP per capita.

Finally, one can aggregate the total employment created by firms from a given population of founders. This creates a simple and intuitive metric, comparing jobs created with the number of workers of a given type.

Corollary 3. *The total number of jobs created by immigrants is greater than the number of immigrants in the workforce if $a_1 > a_0$. The total number of jobs created by immigrants is less than the number of immigrants in the workforce if $a_1 < a_0$.*

Proof. See appendix. □

This corollary is another way of seeing the net labor market implications of immigration. The net labor demand effect (quantity) from immigration, like the net wage effect (price) from immigration, is increasing in immigration if immigrants have an advantageous distribution of entrepreneurial acumen.

4 Empirical Evidence

This section presents empirical evidence about new venture creation in the U.S. economy. We compare native-born and immigrant founded firms. The analysis focuses on the rate of new venture creation and the employment distribution in the firms that results. This section describes the primary data sets analyzed, defines the specific empirical measures, and presents the findings.

4.1 Data

We look at three independent data sets. Our primary analysis uses administrative data to study all new firms in the United States from 2005-2010 ("administrative data"). Our second analysis studies the U.S. Census Bureau's 2012 Survey of Business Owners (SBO) to examine a representative sample of all U.S. businesses and their founders ("representative sample"). Our third analysis, drawing on public sources, studies the Fortune 500 to consider the country's largest firms and their founders ("Fortune 500"). We describe these data sets here, with additional detail in the appendix.

4.1.1 Administrative Data

This data set links the U.S. Census Bureau's Longitudinal Business Database (LBD), population-wide W-2 tax records, and the U.S. Census demographic files (Numident). Business-level information comes from the LBD, which tracks U.S. businesses and their establishments over time. The

LBD includes all non farm private sector firms in the U.S. with at least one employee. Our focus in the LBD is on startups, to examine entrepreneurship by native and immigrant populations. We define startups as *de novo* firms that have no prior activity at any of their establishments. The founding year is the year the firm first appears in the LBD.

To study founders, we integrate employment and demographic records. We use a "founding team" definition, identifying the top three earners (via W-2 records) in each new venture in its founding year (Kerr and Kerr 2017; Azoulay et al. 2020). We then classify each founding team member as either U.S.-born or immigrant, based on the country of birth (via Census Numident). Using W-2 records and Census Numident, we similarly classify all workers in the economy as U.S.-born or immigrant. We will also consider alternative measures of the immigrant population.

To examine employment in the new ventures, we study the firm size five years after founding and compare the firm size distributions for U.S.-born and immigrant founded firms. Given data set availability, our main analysis considers all firms founded from 2005-2010 that survive for five years. We thus analyze employment in 1.02 million new businesses. A benefit of the administrative data is that it covers all new employer firms in the economy. The limitation is that we see employment outcomes only in the early years of the business. This motivates our second data set, which is a representative sample of U.S. business owners, regardless of the business age.

4.1.2   Representative Sample

The Survey of Business Owners (SBO) collects information from a representative sample of U.S. business owners. The SBO is the only source of representative information on economic and demographic characteristics for business owners by gender, ethnicity, race, and veteran status. Included in the collection is information for owners of employer and non-employer businesses. We limit our analysis to employer businesses. The SBO is collected every 5 years. For our exercise we use data from the 2012 SBO.

When looking at business owners, the SBO provides information on the number of owners of the firm but collects detailed information only for the top four owners including their ownership shares, founder status, whether they are native or immigrant, and whether they play an active role in the day-to-day management of the firm. We include all firms regardless of the number of owners

and limit our analysis to businesses with at least one founder amongst the top four owners. Our analysis is strictly limited to the founders (i.e., we exclude investor-owners who did not found the firm). Our analysis sample includes over 200,000 employer firms.

For firm-level employment, we use the survey collected data.[6] We define 10 size classes up to a size class of 10,000 employees. Founding years in our SBO sample range from early 1800 to 2012. While the SBO provides a representative sample, the sample is thin at very large firm sizes. Since a very small number of firms in the economy are extremely large and responsible for substantial employment, we therefore turn additionally to a third data set, the Fortune 500, to explicitly examine this upper tail.

4.1.3 Fortune 500

We further collect data on firms listed in the 2017 edition of the Fortune 500 ranking. For each firm, we capture, whenever possible, the year of incorporation, the name of the founders, and the founders' immigration status. This process is straightforward for many firms, particularly those that were founded in the recent past. For other firms, it is more challenging, since they can be the offspring of many merged entities. Our approach is to trace the "genealogical tree" for each firm to the earliest parent possible and then identify the founders of these parents as the founders of the firm.[7]

Among the Fortune 500, we were able to determine the immigrant status for the founders for 449 firms. The founding years range from 1743 to 2004. Further details regarding the compilation of immigrant status for these firms (including 51 cases where this proved impossible) are provided in the Online Appendix.

4.2 Measures

We define individuals as immigrants if they are born outside the U.S.. For each data set, we count the number of entrepreneurs from each population. We similarly examine the firms' employment distribution, counting firms of a given size from founders in each population. We

---

[6]We replace the survey employment with information from administrative sources when missing. Robustness checks making full use of administrative data do not alter the results.

[7]For instance, American Airlines has two parents, Colonial Air Transport (incorporated in 1926, one native-born founder) and Robertson Aircraft Corporation (incorporated in 1921, two native-born founders).

then normalize these counts by the total size of each population. Together these measures produce the rate of entrepreneurship (E_j/N_j) and the fraction of individuals who start firms of a given size $(n_j\left(l_i^*\right)/N_j)$. In addition to being intuitive ways of characterizing entrepreneurship and labor demand, these empirical measures correspond to the corollaries above.

Firms can have multiple founders, and these founders may mix U.S.-born and immigrant individuals. Allocating firms as either native-founded or immigrant-founded can then be done multiple ways. We therefore consider three different approaches to assigning firms to each sub-population. The first approach, denoted Definition 1, counts the business as an "immigrant firm" if anyone in this founding team is an immigrant. This method is most generous in counting immigrant firms. The second approach, denoted Definition 2, counts the business as an "immigrant firm" only if the "lead" individual is an immigrant. This approach is more conservative. We operationalize the "lead" individual as the highest-paid on the founding team (administrative data) or the owner-worker with the highest ownership share (SBO data).[8] The final approach, denoted Definition 3, uses proportional assignment. It collects all firms in a given employment size range and pools the founding teams among this group of firms. It then allocates these firms proportionally (within the given size range) as native-founded versus immigrant-founded, based on the share of founding team members from each population. These three definitions produce alternative frequency distributions, $n_j\left(l\right)$, and founder counts E_j for each population, acting as robustness checks.

To produce appropriate population normalizations, N_j, for each group, we also have reasonable alternatives. We can consider all individuals who work (e.g., from all W-2 records) in each group, or we can consider broader estimates to account for informal employment and unauthorized immigration. Broader population measures also account for historical immigration levels, which is especially relevant for the Fortune 500 data. In practice, the results appear robust to any plausible estimate of the immigrant population (see Section 5.2).

4.3   Results

The central results are presented in Figures 1, 2, and 3. Each figure shows the normalized firm size distributions for immigrant-founded and native-founded firms. Figure 1 presents the

---

[8]This second approach is not feasible for the Fortune 500 data.

administrative data, Figure 2 presents the representative sample, and Figure 3 presents the Fortune 500 data.

The broad finding is clear, looking across each data set and across different ways of defining immigrant and non-immigrant firms. Namely, immigrant entrepreneurship presents a "right shift" in firm size. We see Zipf's Law appear for both immigrant-founded and native-founded firms, but immigrants tend to start more firms per person of every size. Specifically, at each firm size, the frequency of immigrant-founded firms per immigrant in the population tends to be larger than the frequency of native-founded firms per native-born person in the population. These plotted measures correspond to equation (12).

4.3.1 Administrative Data

Figure 1A presents the administrative data, counting immigrant firms as those with at least one immigrant founder (Definition 1). We see here a parallel right shift for immigrant-founded firms. The rate of entrepreneurship looking at the 2005-2010 period shows that 0.83% of immigrants in the workforce start a firm over this period, compared to 0.46% of native-born individuals in the workforce. Immigrants thus exhibit a 80% higher entrance rate into entrepreneurship. Moreover, immigrants start more firms of every size. They do not just start many small firms, but produce a firm size distribution that looks very similar to native-born entrepreneurs.

Figure 1B considers the administrative data again, but now counts immigrant firms only as those where the highest-paid founding team member is an immigrant (Definition 2). The immigrant firms defined this way are now a strict subset of those pictured in Figure 1A. Conceptually, Definition 1 suggests that each founding team member is necessary for the venture. Definition 2 suggests that only one team member (the highest paid) is necessary for the venture. While the power law for the immigrant-founded firms necessarily moves left compared to Figure 1A, we still see that, at every firm size, immigrant-founded firms outpace the native-founded firms. Now the slope is slightly steeper for immigrant-founded firms, so that they dominate relatively more among small firms, and only slightly among the largest firms.

Figure 1C considers a proportional assignment of immigrant and native-founded firms (Definition 3). The result looks very similar to Definition 2. A small difference is that now, in

13

the very largest size bucket, immigrant-founded firms are slightly outpaced by native-founded firms. Nonetheless, aggregating across firms, the total employment assigned to immigrant-founded firms is greater than the total employment for native-founded firms (as shares of their respective populations). Integrating across the firm size distributions for each population, the total number of jobs created by immigrant-founder firms (per immigrant in the population) is 42% higher than that of native-founder firms (per native-born in the population).

4.3.2 Representative Sample

Figure 2 repeats the analyses of Figure 1, but now using the Survey of Business Owners (SBO). The SBO population includes all firms regardless of age so our analysis is not restricted to young firms. Figure 2A shows the normalized firm size distribution for immigrant vs. native founded firms (Definition 1). As before, we see a right shift in the distribution for immigrant founded firms. The estimated rate of entrepreneurship based on the 2012 set of businesses is 7.25% for immigrants, compared to 4.03% for native-born individuals.[9] Compared to the administrative data, which considers the flow of new firms founded over 2005-2010, this SBO data comes closer to the steady-state looking across the economy's firms. Remarkably, despite measurement differences with the analysis in the previous section, we still find that immigrants exhibit the same 80% higher entrance rate into entrepreneurship when compared to native-born individuals. As before, immigrants disproportionately start more firms across the size distribution.[10]

Figure 2B considers the SBO data again, but now counts immigrant firms as only those whose largest owner-founder is an immigrant (Definition 2). We use the ownership share information to make this determination. Once again we see a right-shift of the firm size distribution for immigrant founded firms but now with a distinct tilt as we move towards the largest size classes. The size distributions converge as we increase in size. This result is again similar to the administrative data, and despite alternative constructs of the lead founder.[11]

---

[9]The underlying immigrant and native population sizes are weighted averages of the respective populations across the founding years of the enterprise. We use data from the Migration Policy Institute for the underlying population counts, available at `https://www.migrationpolicy.org/programs/data-hub/charts/immigrant-population-over-time`.

[10]To comply with disclosure avoidance rules, we suppress data for firms above 10,000 employees as well as specific cells in Figure 2.

[11]The SBO definition, which is based on the largest current ownership share among the founders, might select against immigrant founders to the extent that native-born founders are initially more affluent and have an ownership edge.

Figure 2C considers a proportional assignment of immigrant and native-founded firms (Definition 3). The right-shift of the log size distribution for immigrant founded firms remains. The results are broadly similar as when using the prior definitions, and with continuing similarity with the administrative data findings in Figure 1.

4.3.3 Fortune 500

Figure 3 repeats the prior analysis on the set of firms in the Fortune 500 ranking, using only the first and third definitions. This provides a close look at the very largest firms in the economy and further allows for some historical comparison. In each case, we normalize the count of founders using a count of native and non-native born individuals in the U.S. population in the decade during which the firm was founded.[12] Because the data is more sparse (449 firms with founder information, 96 of which have at least one immigrant founder), we create only three employment bins (using the 2017 employee count for each firm): Firms with less than 30,000 employees, firms with between 30,000 and 100,000 employees, and firms with more than 100,000 employees.

The results of the analysis are depicted in Figures 3A and 3B (corresponding to founder counts according to the first and third definitions, respectively). The patterns observed are consistent with those obtained using the broader administrative and survey data: a right shift for immigrant-founded firms. Among the Fortune 500, the native-born and immigrant slopes are of similar magnitudes. Overall, the Fortune 500 findings indicate that the results extend to the very largest U.S. businesses and to founding behavior over a broader sweep of U.S. business history.[13]

5   Discussion

In this section, we first summarize the results and relate them to the theory. We then consider further interpretations of the findings and related evidence.

---

[12]The source of this information is the Migration Policy Institute tabulation of the U.S. Census Bureau's 2010-2017 American Community Surveys and 1970, 1990, and 2000 Census data, as well as Gibson and Lennon (1999).

[13]In the appendix, we further consider these results focusing on the subset of 123 Fortune 500 firms founded since 1970. This has two advantages. First, the founder information is more comprehensive for this subset, and second the recent subset is more relevant to contemporary immigration outcomes. The results for this recent sample are similar.

5.1 Summary

Overall, immigrants appear highly entrepreneurial. We see a power law in the distribution of firm size for each population, but immigrant entrepreneurship appears right-shifted. Specifically, there are more immigrant-founded firms, per immigrant in the population, at each employment size. This is true recently, where we can look at all new firms in the economy using administrative data. It is also true for firms founded in earlier time periods, including when studying the Survey of Business Owners and the Fortune 500.

Relating to the theory, the higher rate of entrepreneurship (Corollary 1), the right shift in the firm size distribution (Corollary 2), and the total count of jobs created per population member (Corollary 3), all are consistent with immigrants presenting an advantageous distribution of entrepreneurial acumen, compared to native-born individuals. According to Proposition 1, this feature is consistent with increased wages and rising income per capita.

5.2 Unmeasured Immigration

Empirically, the measures depend on estimates of the immigrant and native-born populations. Immigrants who are unaccounted for in government statistics may lead immigrant population measures to be understated. For the SBO and Fortune 500 analyses, our population estimates already explicitly include foreign-born individuals who are residing in the U.S. without authorization (Gibson and Campbell 1999). However, this issue may be particularly germane for populations based on W-2 tax records, and more generally one can assess the scale by which the immigrant population would need to be under-counted for the immigrant and native-born founding behavior to look the same.

As a simple metric to assess the robustness of the findings, we ask how much the underlying immigrant population measure would have to be scaled up so that the aggregate employment in immigrant-founded firms, per immigrant in the population, would decline to the equivalent rate for the native-born. For the W-2 data, where a missing mass of immigrants is potentially more relevant, the immigrant population would have to be scaled up by 42% using Definition 3. Using the Fortune 500 data for founders since 1970, the immigrant population would have to be scaled up

by 64% using Definition 3.[14] To put this required under-count in context, consider that there were approximately 45.6m foreign-born individuals in the U.S. in 2017, which includes approximately 10.5m unauthorized immigrants (Pew Research Center 2018). Higher estimates suggest there are as many as 12 million unauthorized immigrants (Kamarck and Stenglein 2019). Under-counting the total immigrant population by 40-60% would mean that unauthorized immigrants total 30m or more individuals, in comparison to standard estimates of 10.5-12m. There is no evidence that immigration could be understated by anything close to this magnitude. Overall, immigration on net appears to be a net job creator in the U.S. economy when including unauthorized immigrants.

5.3 Wages

The conceptual framework has emphasized heterogeneity in entrepreneurial acumen and resulting outcomes in the firm size distribution. More generally, one can look at labor market impacts of founders not just through the employment in firms they create, but also in the wages these firms pay. Specifically, one might wonder whether immigrant-founders, although they create a large number of jobs, perhaps do not create high-paying jobs.

The appendix and Table A.1 use W-2 tax records to estimate worker-level wage regressions, comparing the wages paid in immigrant-founded versus native-founded firms in the administrative data. In a bivariate regression, workers in immigrant-founded firms receive 4.1% higher wages on average. Controlling for founding year and county, however, the wages become identical for workers in immigrant and native founded firms. Additionally controlling for sector as well as worker characteristics, including age, gender, and immigrant status, the wage differences continue to shift somewhat and can flip sign. With all the controls we find that the wages are similar, with a slightly higher wage (0.7%) in immigrant-founded firms. Overall, these findings suggest that immigrant-founders not only are substantial job creators but also do not appear to create lower paying jobs.

---

[14]The required scaling is similar using Definition 2, where Definition 2 is available. The required scaling is larger using Definition 1.

5.4   Small Businesses

The distributional right shift for immigrant-founded firms, while appearing across firm sizes, can present a larger shift among small businesses. This feature is seen more with Definitions 2 or 3, where there is a steeper slope in the firm size distribution for immigrant-founded firms. As one interpretation, the relatively large set of small businesses may be consistent with a "push" mechanism for immigrant entrepreneurs (Light and Roach 1996; Fairlie and Lofstrom 2010). In this view, immigrants may have relatively limited economic opportunities as workers, for example due to limited social networks or discrimination, which can manifest itself as trouble finding jobs or being paid less when they do. Such forces may push immigrants to start businesses at higher rates. To the extent these individuals may have somewhat less business acumen, one may also expect the resulting firms to remain relatively small. The preponderance of small businesses is also consistent with ongoing limited opportunities as owners due to similar forces that limit labor prospects. Related, the tilt in moving from Definition 1 to Definition 2 is consistent with immigrant founders having fewer initial financial resources and a somewhat lower ownership share in the new firm. Unpacking potential small-business-oriented mechanisms are important areas for further work. At the same time, our central finding is a more systematic shift in the firm-size distribution, suggesting the broader entrepreneurial orientation and success of immigrants in starting firms, including high-growth firms.

5.5   Technology Businesses

A different literature emphasizes that immigrants are disproportionately likely to hold STEM degrees (Kerr and Kerr 2020) and play large roles in major entrepreneurial ecosystems like Silicon Valley (Saxenian 2002). To study the technological and inventive orientation of immigrant versus non-immigrant startups, we further consider patenting behavior. This analysis links the corpus of U.S. patents to each firm in the administrative data, studying all firms founded over the 2005-2010 period. Figure 4 presents the results. Overall, firms with an immigrant founder are roughly 35% more likely to have a patent than firms with no immigrant founders. Studying firms by size group, immigrants are more likely to have patents at all sizes and especially at larger sizes.[15]

---

[15]These results are largely consistent with Brown et al. (2020) who find higher rates of innovation linked to immigrant-founded startups across several dimensions including R&D and trademarks.

To the extent that inventive firms bring productivity gains beyond the bounds of the firm, entrepreneurship can play additional welfare roles. Large literatures find substantial spillovers and high social returns from innovation investments (e.g., Hall et al. 2010, Bloom et al. 2013, Jones and Summers 2020) and emphasize that technology advances play critical roles in driving rising standards of living (e.g., Solow 1956, Mokyr 1990, Cutler et al. 2006). Conceptually, the model in Section 3 emphasized a general equilibrium allocation without innovative spillovers. Adding productivity spillovers from inventive firms leads to the intuitive result that immigrant entrepreneurship can further enhance productivity, wages, and per-capita income in the economy.[16] In this context, pushing ahead the technology frontier of the economy acts as an additional benefit.

5.6 Immigration and Selection

Overall, the picture is a rightward shift in the firm size distribution. While there is a potentially rich set of underlying mechanisms, the findings appear consistent overall with immigrants being positively selected on entrepreneurial acumen. Various forces may explain this. Low-ability individuals may face difficulties migrating and immigration policy deploys many visa classes that select on high ability (Chiswick 1999). The empirical literature on migration has come to mixed conclusions but often suggests that immigrants are positively selected on ability from their home communities (McKenzie et al. 2010, Hendricks and Schoellman 2018). More broadly, the act of migration itself, with its attendant embrace of risk and the unknown, may suggest an entrepreneurial orientation; for example, the historical literature in the United States emphasizes a "frontier" spirit, associated with adventurous migrants and "a practical, inventive turn of mind" (Turner 1921), and contemporary literature has found that migrants are less risk averse (Jaeger et al. 2010). These various perspectives may all suggest a right-shift in entrepreneurial acumen among immigrants.

6 Conclusion

This paper has studied the relative roles of immigrant and native-born individuals in new venture formation in the United States. Using administrative data, a representative sample, and Fortune 500 data, we present new findings on the size of firms these different founder populations create. Across all three data sets, we find that immigrants present a "right shift" in new venture

---

[16]A formal model that includes such spillovers is available from the authors upon request.

formation, where immigrants start more firms of each size per member of their population. A simple theoretical framework provides intuition for thinking about these roles and helps make the measures precise.

Overall, the entrepreneurial lens suggests that immigrants appear to play a relatively strong role in expanding labor demand relative to labor supply, compared to the native-born population. These findings can help resolve the tension between labor supply oriented analyses (e.g., Isaac 1947, Borjas 1994, Dustmann et al. 2016), where immigrants are seen to compete with local workers and depress wages, and natural experiments that often show more positive economic results of immigration for native born workers (e.g., Card 1990, Hunt 1992, Friedberg 2001). At the same time, immigrants can play broader economic roles than examined in this paper, and additional theoretical and empirical approaches can frame further dimensions. For example, immigration can have fiscal implications (e.g., Storesletten 2000), implications for the emigrant countries (e.g., Giuliano and Ruiz-Arranz 2009, Docquier and Rapoport 2012), and political economy implications (e.g., Tabellini 2020). Embracing these dimensions in further research can help develop an increasingly full picture of migration and its effects.

References

Altonji, Joseph G., and David Card. 1991. "The Effects of Immigration on the Labor Market Outcomes of Less-skilled Natives." In John M. Abowd, and Richard B. Freeman (Eds.), *Immigration, Trade, and the Labor Market*, pp. 201-234. Chicago, IL: University of Chicago Press.

Axtell, Robert L. 2001. "Zipf Distribution of U.S. Firm Sizes." *Science* **293**(5536): 1818-1820.

Azoulay, Pierre, Benjamin F. Jones, J. Daniel Kim, and Javier Miranda. 2020. "Age and High-Growth Entrepreneurship." *American Economic Review: Insights* **2**(1): 65-82.

Bernstein, Shai, Rebecca Diamond, Timothy James McQuade, and Beatriz Pousada. 2019. "The Contribution of High-Skilled Immigrants to Innovation in the United States." Stanford Graduate School of Business Working Paper No. 3748.

Bloom, Nick, Mark Schankerman, and John van Reenen. 2013. "Identifying Technology Spillovers and Product Market Rivalry." *Econometrica* **81**(4): 1347-1393.

Borjas, George J. 1994. "The Economics of Immigration." *Journal of Economic Literature* **32**(4): 1667-1717.

Borjas, George J. 2003. "The Labor Demand Curve *Is* Downward Sloping: Reexamining the Impact of Immigration on the Labor Market." *Quarterly Journal of Economics* **118**(4): 1335-1374.

Briggs, Jr., Vernon M. 2001. *Immigration and American Unionism.* Ithaca, NY: Cornell University Press.

Brown, J. David, John S. Earle, Mee Jung Kim, and Kyung Min Lee. 2020. "Immigrant Entrepreneurs and Innovation in the US High-Tech Sector." In Ina Ganguli, Shulamit Kahn, and Megan MacGarvie (Eds.), *The Roles of Immigrants and Foreign Students in US Science, Innovation, and Entrepreneurship*, pp. 149-171. Chicago, IL: University of Chicago Press.

Card, David. 1990. "The Impact of the Mariel Boatlift on the Miami Labor Market." *Industrial and Labor Relations Review* **43**(2): 245-257.

Chiswick, Barry. 1999. "Are Immigrants Favorably Self-selected?" *American Economic Review* **89**(2): 1999.

Constant, Amelie, and Klaus F. Zimmermann. 2006. "The Making of Entrepreneurs in Germany: Are Native Men and Immigrants Alike?" *Small Business Economics* **26**(3): 279-300.

Cutler, David, Angus Deaton, and Adriana Lleras-Muney. 2006. "The Determinants of Mortality." *Journal of Economic Perspectives* **20**(3): 97-120.

Docquier, Frédéric, and Hillel Rapoport. 2012. "Globalization, Brain Drain, and Development." *Journal of Economic Literature* **50**(3): 681-730.

Dustmann, Christian, Uta Schönberg, and Jan Stuhler. 2016. "The Impact of Immigration: Why Do Studies Reach Such Different Results?" *Journal of Economic Perspectives* **30**(4): 31-56.

Fairlie, Robert W., and Magnus Lofstrom. 2015. "Immigration and Entrepreneurship." In Barry R. Chiswick, and Paul W. Miller (Eds.), *Handbook of the Economics of International Migration*, pp. 877-911. Amsterdam: Elsevier: North-Holland.

Friedberg, Rachel M. 2001. "The Impact of Mass Migration on the Israeli Labor Market." *Quarterly Journal of Economics* **116**(4): 1373-1408.

Gibson, Campbell, and Emily Lennon. 1999. "Historical Census Statistics on the Foreign-Born Population of the United States: 1850 to 1990." Working Paper No. 29, U.S. Bureau of the Census, Population Division.

Giuliano, Paola, and Marta Ruiz-Arranz. 2009. "Remittances, Financial Development, and Growth." *Journal of Development Economics* **90**(1): 144-152.

Hall, Bronwyn H., Jacques Mairesse, and Pierre Mohnen. 2010. "Measuring the Returns to R&D." In Bronwyn H. Hall, and Nathan Rosenberg (Eds.), *Handbook of the Economics of Innovation*, pp. 1033-1082. New York: North-Holland.

Hendricks, Lutz, and Todd Schoellman. 2018. "Human Capital and Development Accounting: New Evidence from Wage Gains at Migration." *Quarterly Journal of Economics* **133**(2): 665-700.

Hunt, Jennifer. 1992. "The Impact of the 1962 Repatriates from Algeria on the French Labor Market." *Industrial and Labor Relations Review* **45**(3): 556-572.

Isaac, Julius. 1947. *Economics of Migration*. New York: Oxford University Press.

Jaeger, David A., Thomas Dohmen, Armin Falk, David Huffman, Uwe Sunde, and Holger Bonin. 2010. "Direct Evidence on Risk Attitudes and Migration." *Review of Economics and Statistics* **92**(3): 684-689.

Jones, Benjamin, and Lawrence H. Summers. 2020. "The Social Returns to Innovation Investments." Working Paper, Northwestern University.

Kamarck, Elaine, and Christine Stenglein. 2019. "How many undocumented immigrants are in the United States and who are they?" *Brookings Policy 2020*. Available at `https://www.brookings.edu/policy2020/votervital/how-many-undocumented-immigrants-are-in-the-united-states-and-who-are-they/`.

Kerr, Sari Pekkala, and William Kerr. 2017. "Immigrant Entrepreneurship." In John Haltiwanger, Erik Hurst, Javier Miranda, and Antoinette Schoar (Eds.), *Measuring Entrepreneurial Businesses: Current Knowledge and Challenges*, pp. 187-249. Chicago, IL: University of Chicago Press.

Kerr, Sari Pekkala, and William Kerr. 2020. "Immigrant Entrepreneurship in America: Evidence from the Survey of Business Owners 2007 & 2012." *Research Policy* **49**(3): 103918.

Light, Ivan, and Elizabeth Roach. 1996. "Self-Employment: Mobility Ladder or Economic Lifeboat?" In Roger Waldinger, and Mehdi Bozorgmehr (Eds.), *Ethnic Los Angeles*, pp. 193-214. New York: Sage.

Lucas, Robert. 1978. "On the Size Distribution of Business Firms." *Bell Journal of Economics* **9**(2): 508-523.

McKenzie, David, Steven Stillman, and John Gibson. 2010. "How Important is Selection? Experimental vs. Non-Experimental Measures of the Income Gains from Migration." *Journal of the European Economic Association* **8**(4): 913-945.

Mokyr, Joel. 1990. *The Lever of Riches: Technological Creativity and Economic Progress*. New York: Oxford University Press.

Passel, Jeffrey S., and D'Vera Cohn. 2018. "U.S. Unauthorized Immigrant Total Dips to Lowest Level in a Decade." Pew Research Center. Available at `https://www.pewresearch.org/`

hispanic/2018/ 11/27/u-s-unauthorized-immigrant-total-dips-to-lowest-level-in-a-decade/.

Reder, Melvin W. 1963. "The Economic Consequences of Increased immigration." *Review of Economics and Statistics* **45**(3): 221-230.

Saxenian, AnnaLee. 2002. "Silicon Valley's New Immigrant High-Growth Entrepreneurs." *Economic Development Quarterly* **16**(1): 20-31.

Sequeira, Sandra, Nathan Nunn, and Nancy Qian. 2020. "Immigrants and the Making of America." *The Review of Economic Studies* **87**(1): 382-419.

Solow, Robert M. 1956. "A Contribution to the Theory of Economic Growth." *Quarterly Journal of Economics* **70**(1): 65-94.

Storesletten, Kjetil. 2000. "Sustaining Fiscal Policy Through Immigration." *Journal of Political Economy* **108**(2): 300-323.

Tabellini, Marco. 2020. "Gifts of the Immigrants, Woes of the Natives: Lessons from the Age of Mass Migration." *The Review of Economic Studies* **87**(1): 454-486.

Turner, Frederick Jackson. 1921. *The Frontier in American History.* New York: Henry Holt & Co.

Figure 1
Immigrant and Native-Born Entrepreneurship:
Firm Size Distributions using Administrative Data
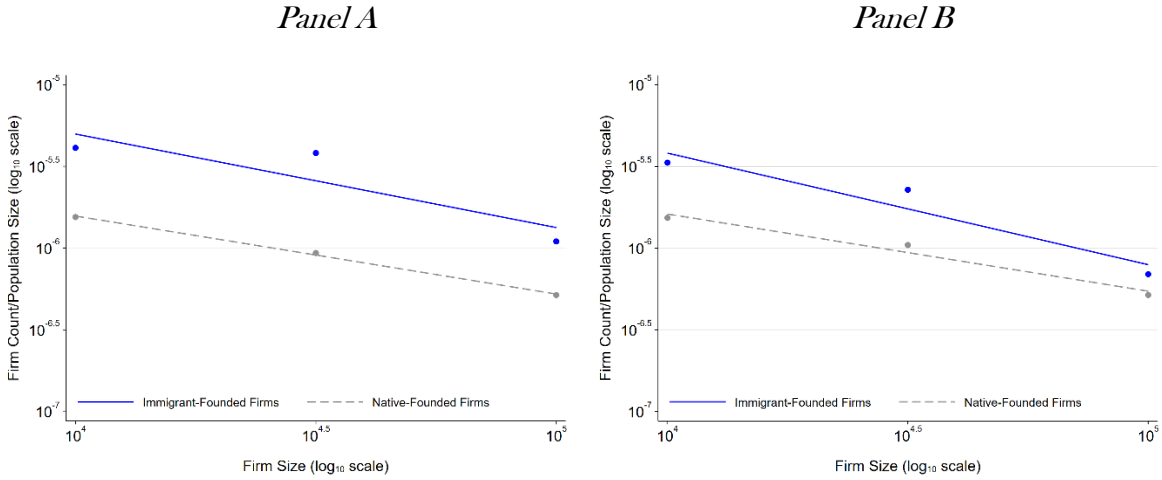
Panel A
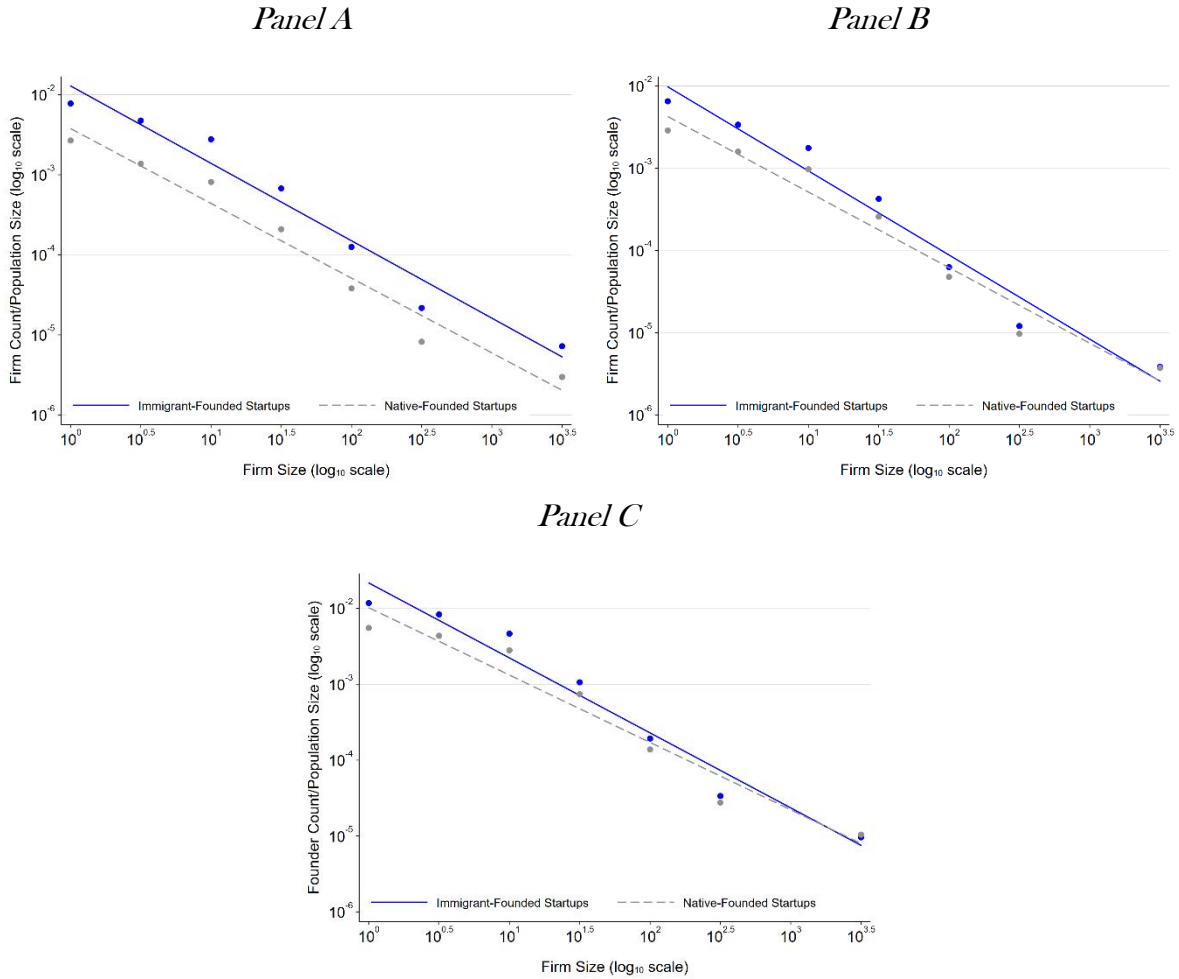
Panel B

Panel C

<u>Notes</u>: Each panel consider the firm size distribution, distinguishing between immigrant-founded and native-founded firms, for all U.S. firms founded in the 2005-2010 period. The x-axis is the log of firm size measured as total employment in the firm five years after founding. The y-axis is the log count of firms of a given size, with the count normalized by the number of workers from the relevant population (immigrant or native born). The plotted measures correspond to equation (12). Panel A counts a firm as immigrant-founded if any of the founding team members are immigrants (Definition 1 in the text). Panel B counts a firm as immigrant-founded only if the highest paid member of the founding team is an immigrant (Definition 2). Panel C assigns firms to immigrant and non-immigrant proportionally based on the mix of immigrant and native-born individuals in the founding teams (Definition 3).

24

Figure 2
Immigrant and Native-Born Entrepreneurship:
Firm Size Distributions using Survey of Business Owners
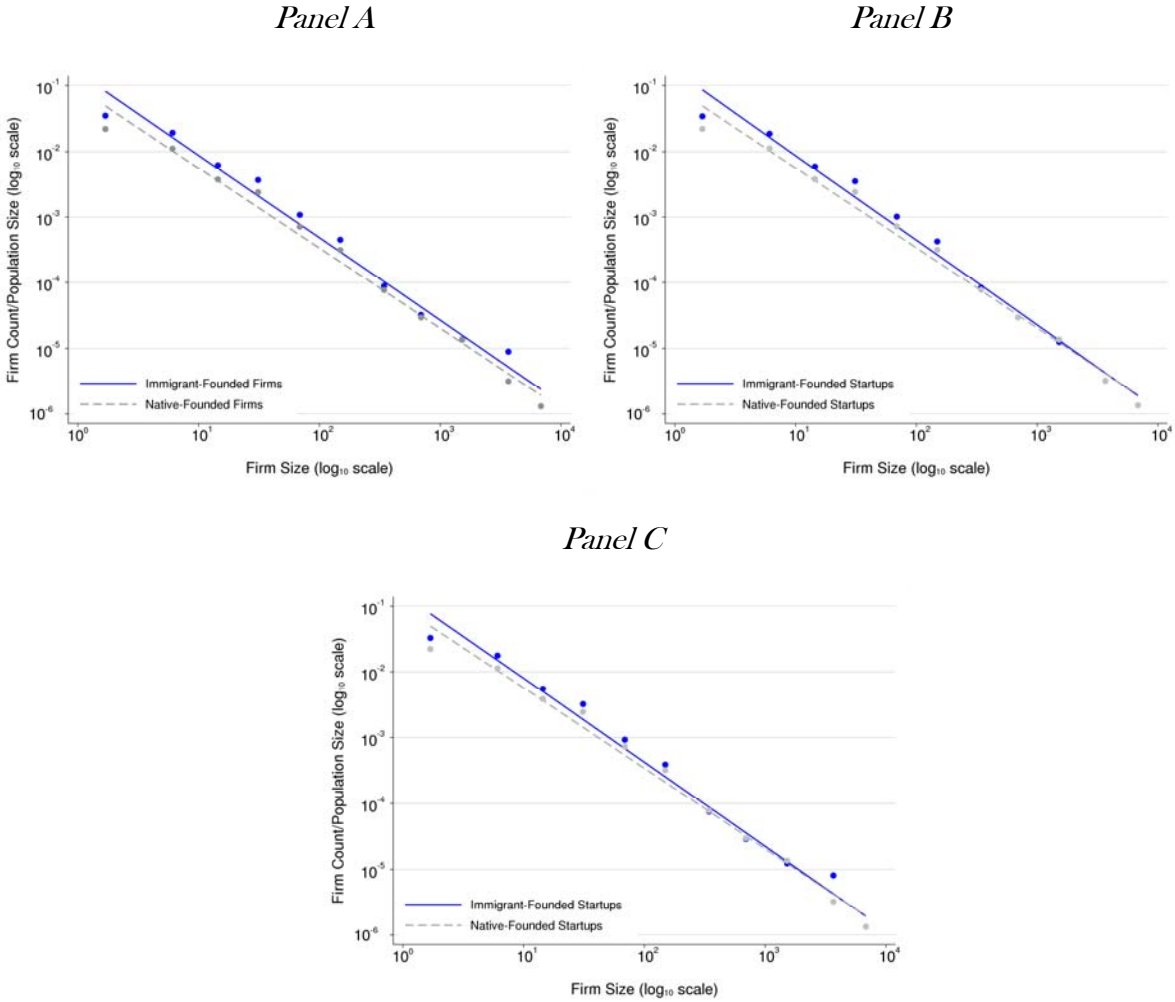
Panel A　　　　　　　　　　　　　　　　　　*Panel B*


Panel C

Notes: Each panel consider the firm size distribution, distinguishing between immigrant-founded and native-founded firms, using a representative sample of U.S. firms from the 2012 Survey of Business Owners. The x-axis is the log of firm size measured as current total employment in the firm. The y-axis is the log count of firms of a given size, with the count normalized by the population size of the relevant group (immigrant or native born). The population measure is an average of the immigrant or native-born population size in the year of founding, weighted by the number of firms founded in that year. The plotted measures correspond to equation (12). Panel A counts a firm as immigrant-founded if any of the owner-founders are immigrants (Definition 1 in the text). Panel B counts a firm as immigrant-founded only if the owner-founder with the highest current ownership share is an immigrant (Definition 2). Panel C assigns firms to immigrant and non-immigrant founded proportionally based on the mix of immigrant and native-born individuals among the owner-founders (Definition 3).

Figure 3

Figure 3
Immigrant and Native-Born Entrepreneurship:
Firm Size Distributions using the Fortune 500

Panel A   *Panel B*

<u>Notes</u>: Each panel consider the firm size distribution, distinguishing between immigrant-founded and native-founded firms. The x-axis is the log of firm size measured as current total employment in the firm, using the 2017 Fortune 500. The y-axis is the log count of firms of a given size, with the count normalized by the population size for the relevant group (immigrant or native-born). The population measure is an average of the immigrant or native-born population in the decade of founding, weighted by the number of firms founded in that decade. The plotted measures correspond to equation (12). Panel A counts a firm as immigrant-founded if any of the founders are immigrants (Definition 1 in the text). Panel B assigns firms to immigrant and non-immigrant proportionally based on the mix of immigrant and native-born founders of the initial business (Definition 3). Definition 2 is not available for the Fortune 500, as discussed in text.

Figure 4
Technology-Based Immigrant and Native-Born Entrepreneurship
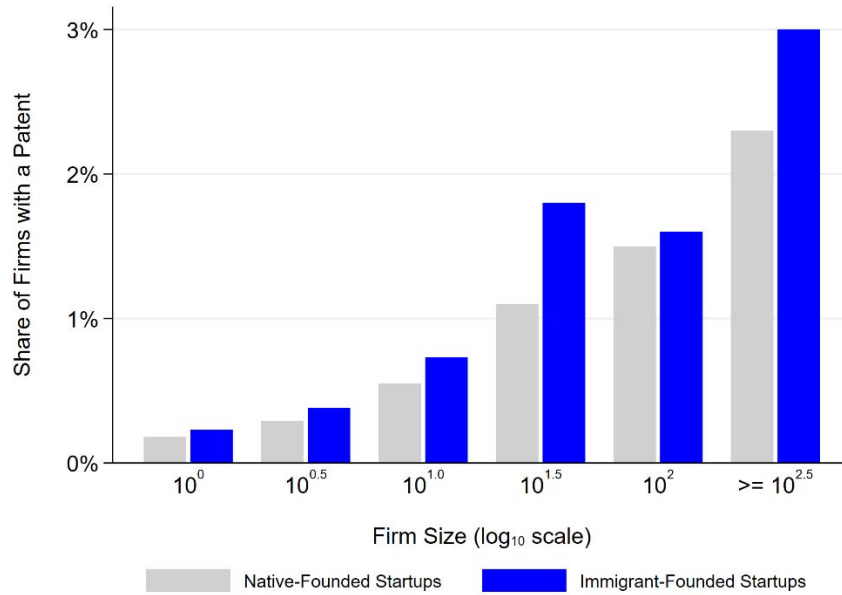
Notes: Using W-2 and LBD data combined with patenting records from the USPTO, this figure shows the share of firms in each firm size bin that own at least one patent, distinguishing between native-founded versus immigrant-founded startups, for all firms in the US between 2005 and 2010. Immigrant-founded startups are identified using Definition 1, which equals 1 if at least one of the founders are foreign-born. Firms are grouped into six bins according to the number of employees five years after founding.

Supplementary Online Material

Appendix A: Data

U.S. Census Data

In this appendix, we describe the various data sets used in this study. Many of the data sets are Census-based products which are available to researchers through Census-approved projects and accessible through Federal Statistical Research Data Centers (FSRDC). Form W-2 data are currently accessible only by Census employees who have been granted access through approved internal projects.

The Longitudinal Business Database (LBD). The LBD is a panel dataset of all establishments in the U.S. with at least one paid employee. This dataset begins in 1976 and currently runs through 2015. The coverage includes all industries in the private non-farm sector and every state in the U.S.. The LBD is sourced from administrative income and payroll filings and enhanced with other Census data sets, including the Economic Census and the Company Organization Survey. The LBD contains information on the firm size, firm age, location, payroll, legal form of entity, and other characteristics of the establishment.

Form W-2. Our annual individual earnings information are sourced from Form W-2, which is a tax form used to report income paid to employees for their services rendered. Employers are linked to the LBD based on their employer identification numbers (EIN). The W-2 database in the Census begins in 2005 and covers through 2016. Key variables in Form W-2 include income, social security taxes, and Medicare taxes.

The Survey of Business Owners (SBO). Information on the immigrant vs native-born nature of entrepreneurs is obtained from the 2012 Survey of Business Owners (SBO). The SBO collects information about characteristics of the businesses and their owners from a representative sample of firms in the U.S. The random sample of businesses was selected from a list of all firms operating during 2012 with receipts of $1,000$ dollars or more. The SBO universe was stratified by state, industry, owner characteristics, and whether the company had paid employees in 2012. Large companies were selected with certainty. The remaining universe was subjected to stratified systematic random sampling. Each firm selected into the sample was asked the percentage of ownership, gender, ethnicity, race, and veteran status for up to four persons owning the largest percentages in the business. The final sample includes over 200,000 employer businesses in the SBO. Each firm in the SBO sample is assigned a weight equal to the reciprocal of the firm's probability of selection. Certainty cases are given a weight of one. Sample weights are used in the calculation of the results reported in the paper as frequency weights to return the population totals.

Census Numerical Identification System File (NUMIDENT). In order to define immigrant entrepreneurs, we use foreign-born status of individuals in the NUMIDENT. This Census database is originally sourced from the Social Security Administration (SSA) applications for Social Security Numbers (Form SS-5). Other person-level characteristics are contained in the NUMIDENT including gender, ethnicity, and date of birth.

The Patent Longitudinal Business Database Crosswalk (LPBD). The LPBD links patents data from the U.S. Patents and Trademark Office (USPTO) to firms in the LBD. This database begins in 2000 and extends to 2015. Though both application and grant years of the patent are observed, only granted patents are included in this sample. Other key variables include assignee location and type.

Fortune 500 Data

We collected founder and founding information for the firms listed in the 2017 edition of the Fortune 500 ranking. For each firm, we capture, whenever possible, the year of incorporation, the name of the founder, and his/her country of birth. This data collection builds on earlier efforts by the New American Economy Research Fund (2011, 2018)

and the Center for American Entrepreneurship (2017).[i] We extend their analysis by including all founders for these firms, whether U.S.-born or immigrants.

This process is straightforward for many firms, particularly those that were founded in the recent past.[ii] For others, it is more challenging, since they are the offspring of many merged entities. Our approach is to walk back the genealogical tree of each firm to the earliest parent possible, and then to identify the founders of these parents. A firm will therefore have potentially many founders because it has multiple parents.

There are also particular cases where we do not include the firm. For some firms (particularly railroads and power utilities), there are very many mergers and it is not possible to trace the founders effectively. Further, web searches and the *Who's Who* occasionally do not enable us to ascertain the place of birth of any of the firm's original founders. If we cannot determine immigration status for any founder, the firm is dropped from the analysis. Separately, some firms listed in the Fortune 500 were not created through acts of entrepreneurship, but rather by government fiat (Fannie Mae is such an example; Delek U.S. holdings, the state-owned Israel oil company is another one). We exclude these firms from the analysis since they cannot be said to have founders in the traditional sense.[iii] Overall, the sample includes 449 firms and 730 founders for whom we can determine country of birth.

Post-1970 sample. As an additional check on the Fortune 500 analysis, we also consider firms founded since 1970. This includes 117 firms (and 223 founders with country of birth information) in the Fortune 500 ranking. We additionally focus on this time period for two reasons. First, the ability to identify founders —and to ascertain their country of birth —is greater when focusing on firms founded in the more recent past. Second, the recent subset may be most relevant to understanding links between entrepreneurship and immigration in a contemporary setting.

Population Data

The firm size distributions and rate of entrepreneurship measures are normalized by the population size of the relevant group (U.S.-born and immigrant individuals). To ascertain these population sizes we use two different methods, depending on the data source. We also consider robustness tests.

For the administrative data, we use the underlying, complete population of W-2 workers. All individuals with W-2's in the U.S. economy are matched to Census NUMIDENT to code U.S. born and foreign-born workers. This analysis covers these populations of workers from 2005-2010 to match with the founding years we study.

For the SBO data and the Fortune 500 data, the founding years of the firms extend back over many decades. For historical population estimates, we rely on numbers contained in U.S. censuses and collated by the Migration Policy Institute.[iv] This data provides estimates of the immigrant population for each decade from 1850-2010 (and annual estimates thereafter). These data explicitly include estimates of the unauthorized immigrant population.

Population weights. Since the immigrant population share changes over time, and the SBO and Fortune 500 data include a wide range of founding years, we calculate a weighted population over the relevant distribution of founding years. Specifically, for the firm size distributions, in each size bin $\times$ immigration status cell, we normalize the count of firms by the group's population. This population is the weighted averaged across the distribution of founding years of the firms in that size bin.

Unauthorized immigrant population. The census population data in each time period includes all individuals present in the U.S., regardless of citizenship or legal immigration status. In practice, demographers have long recognized that undocumented immigrants are less likely to participate in census surveys, a source of "coverage error" that is then corrected for in these population counts (Van Hook and Bachmeier 2013; Pew Research Center 2018). Of

---

[i]See `https://startupsusa.org/fortune500/` and `https://research.newamericaneconomy.org/report/new-american-fortune-500-in-2018-the-entrepreneurial-legacy-of-immigrants-and-their-children/`.

[ii]Think for example of Hewlett-Packard: incorporated in 1939, with two founders, both native born. Or Google: incorporated in 1998, with two founders, one native-born, the other an immigrant.

[iii]A related example is that of Targa Resources. Warburg Pincus engineered a merger to create this firm in 2003, but it would be wrong to list as its founder Eric Warburg, who created the investment bank back in 1900.

[iv]Available at `https://www.migrationpolicy.org/programs/data-hub/charts/immigrant-population-over-time`. Immigration figures for decades prior to 1970 stem from the work of Gibson and Campbell (1999).

note, disagreements regarding estimates of the undocumented immigrant population occur within a relatively narrow range.[v]

While there is no obvious bias in these population estimates, one may nonetheless consider how sensitive the results in the paper could be to any under-count of the immigrant population. Specifically, how large would an under-count need to be for the right-shift observed between the firm size distribution of immigrants and non-immigrants to vanish?

To analyze this question, we proceed as follows. We define a fixed percentage α by which the immigration population is under-counted in each time period. We then compute the critical value α_0 that would make the total employment in the immigrant-founded firms as a proportion of the immigrant population equal to the total employment in the native-founded firms as a proportion of the native population. This percentage scaling of the immigrant population, α_0, is reported in Section 5.2.

Appendix B: Additional Results

Fortune 500 firms, Post 1970

As an additional view of the Fortune 500 data, Appendix Figure B1 repeats Figure 3 but now focusing on the (2017) Fortune 500 firms that are founded from 1970 onward. This subset includes 123 firms. As can be seen, the results are similar.

Wages

The empirical results investigate employment and the firm size distribution for native-founded and immigrant-founded firms. The empirical analysis in turn follows our conceptual framework, where individuals have heterogeneous entrepreneurial acumen and, for simplicity, homogeneous labor.

Of additional interest may be the wages for the jobs that these founders create and how these wages compare between immigrant and native-founded firms. The administrative data, with which we have integrated the W-2 records for every individual working in these firms, provides an additional opportunity to examine wages in a systematic fashion. We run OLS regressions of the form

$$\log(w_i) = \beta ImmigrantFounded_f + \gamma X_i + \theta Z_f + \epsilon_i$$

where w_i is the individual worker i's annual W-2 earnings from employer f, X_i is a vector of the individual worker characteristics, and Z_f is a vector of the firm's characteristics. Individual worker characteristics include fixed effects for age and indicators for gender and for being foreign-born. Firm characteristics include fixed effects for founding year, fixed effects for county, and fixed effects for sector using NAICS 4-digit industry codes. Results are presented in Table B1 and are discussed in Section 5.3.

Patenting

Figure 4 presents the patenting rate by firm size, comparing immigrant-founded and native-founded firms, in the administrative data. This analysis uses Definition 1 for defining immigrant firms. That is, we consider firms as immigrant-founded if at least 1 of the founders is an immigrant. Appendix Figure B2 repeats the analysis but uses Definition 2 instead. In this definition, the firm is an immigrant-founded firm only if the highest-wage individual in the founding team is an immigrant. (By construction, now the native-founded firms include some founding teams that include immigrants.) Under this definition, immigrant-founded teams still have a higher rate of patenting in each size bucket, although the difference is not as large.

---

[v]Despite using slightly different data and assumptions, estimates from the Pew Research Fund, the Department of Homeland Security, and the Center for Migration Studies have never differed by more than 1 million people, less than 10% of the total unauthorized population.

Appendix C: Proofs

Proof of Proposition 1

This proof proceeds in two steps. First, we consider how a shift in $f(a)$ influences the equilibrium a^*. Second, we consider how other equilibrium outcomes shift. To clarify the comparative statics, we will write $f(a|\theta)$, where θ is a parameter that affects the distribution of entrepreneurial talent. In particular, θ can be the fraction of immigrants in the economy. Equilibrium outcomes will in general be functions of θ.

Comparative Statics on a^*

To begin, we look at a^*. There are two key equations to develop the relevant comparative statics. The first equation comes from the free entry condition to entrepreneurship, defining a threshold value of entrepreneurial talent at which people start firms as opposed to being workers. This relationship is (4)

$$w = a^* (1 - \beta)^{1-\beta} \beta^\beta$$

indicating that there is a monotonically increasing relationship $w(a^*)$.

The second equation comes from the resource constraint, which tells us that the number of entrepreneurs (E) and the number of workers (L) must add up to the total number of people, N. Given the distribution $f(a|\theta)$, we can then write the mapping between the threshold value for founding a firm, a^*, and the number of entrepreneurs as in (5)

$$\frac{E^*}{N} = \int_{a^*}^{\infty} f(a_i|\theta)\, da_i$$

where we are using the fact that anyone with $a_i \geq a^*$ will start a firm (because income as a founder then exceeds income as a worker).

Similarly, the number of workers at a given firm is $l_i^* = \left(\frac{\beta a_i}{w}\right)^{\frac{1}{1-\beta}}$, where a_i is the acumen of the founder. Integrating across all firms we have the total labor force L^*. We can then write that the resource constraint, $\frac{E^*}{N} + \frac{L^*}{N} = 1$, as

$$\int_{a^*}^{\infty} \left[1 + \left(\frac{\beta}{w}\right)^{\frac{1}{1-\beta}} a_i^{\frac{1}{1-\beta}}\right] f(a_i|\theta)\, da_i = 1 \tag{13}$$

This gives us our second function for $w(a^*)$. Using the entrepreneurial entry condition, (4), we can then rewrite this resource constraint to eliminate the wage and put everything in terms of a^*. Namely,

$$\int_{a^*}^{\infty} \left[1 + \frac{\beta}{1-\beta} \left(\frac{a_i}{a^*}\right)^{\frac{1}{1-\beta}}\right] f(a_i|\theta)\, da_i = 1 \tag{14}$$

To interpret this expression, note that we are counting up the number of people at each firm, which must sum to all the people in the economy. We have divided by N so we are counting people in terms of fractions of the population. The term in square brackets is the number of people associated with a given firm. The 1 in square brackets is the entrepreneur – every firm has 1 entrepreneur. The second term in square brackets, $\frac{\beta}{1-\beta} \left(\frac{a_i}{a^*}\right)^{\frac{1}{1-\beta}}$, is the number of workers at that firm, which is increasing in the acumen of the entrepreneur. The $f(a_i)$ then gives the mass of the founder population associated with that firm.

The core result is then seen directly. By inspection, the term in square brackets is strictly positive. Therefore, if you increase the mass of $f(a_i)$ for all points $a_i > a^*$, then the value of the integral would rise. The only way for the integral value to remain constant is therefore for a^* to rise. And if a^* rises, then the wage has to rise, per (4).

More formally, one can takes the comparative statics for $a^*(\theta)$ using Leibniz's Rule. Differentiating (14) with respect to θ, we find that

$$a^{*\prime}(\theta) = \frac{1-\beta}{f(a^*(\theta)) + \frac{1}{a^*(\theta)}\frac{L^*}{N}} \int_{a^*(\theta)}^{\infty} \left[1 + \frac{\beta}{1-\beta}\left(\frac{a_i}{a^*(\theta)}\right)^{\frac{1}{1-\beta}} \right] \frac{df(a_i|\theta)}{d\theta} da_i$$

By inspection, the sign of $a^{*\prime}(\theta)$ depends on the sign of the integral. One can then generate necessary and sufficient conditions for the comparative statics by evaluating the integral for known probability distributions and shifts in these distribution. However, since the term in square brackets is strictly positive, we can also develop simple sufficient conditions that generalize across $f(a)$. In particular, consider the comparative static on the share of immigrants in the economy, defined as $\theta = n_1 = N_1/N$. The population distribution of entrepreneurial acumen is $f(a_i) = (1-n_1)f_0(a_i) + n_1 f_1(a_i)$ and thus

$$\frac{df(a_i|n_1)}{dn_1} = f_1(a_i) - f_0(a_i)$$

It then follows that

$$
\begin{aligned}
a^{*\prime}(n_1) &> & 0 \text{ if } f_1(a_i) > f_0(a_i) \text{ for all } a_i \geq a^* \\
a^{*\prime}(n_1) &= & 0 \text{ if } f_1(a_i) = f_0(a_i) \text{ for all } a_i \geq a^* \\
a^{*\prime}(n_1) &< & 0 \text{ if } f_1(a_i) < f_0(a_i) \text{ for all } a_i \geq a^*
\end{aligned}
$$

which correspond to the three cases in the text and the first part of Proposition 1, as was to be shown.

The comparative statics on other equilibrium quantities are then as follows.

Comparative Statics on w^*

From (4), the equilibrium wage w^* is monotonically increasing in a^*. Hence, the effect of increased immigration on equilibrium wages has the same sign as the comparative statics for a^*, as was to be shown.

Comparative Statics on Y^*/N

From the income side, we can write GDP per capita, $y = Y/N$, as

$$y = \int_{a_m}^{a^*} wf(a_i)da_i + \int_{a^*}^{\infty} \pi_i f(a_i)da_i$$

Using Leibniz's rule, we have

$$
\begin{aligned}
\frac{\partial y}{\partial \theta} &= \frac{\partial w}{\partial \theta}\int_{a_m}^{a^*} f(a_i)da_i + w\int_{a_m}^{a^*}\frac{\partial f(a_i)}{\partial \theta}da_i + wf(a^*)\frac{\partial a^*}{\partial \theta} \\
&+ \int_{a^*}^{\infty}\frac{\partial \pi_i}{\partial \theta}f(a_i)da_i + \int_{a^*}^{\infty}\pi_i\frac{\partial f(a_i)}{\partial \theta}da_i - \pi(a^*)f(a^*)\frac{\partial a^*}{\partial \theta}
\end{aligned}
$$

Noting that $w = \pi(a^*)$, the third and sixth terms cancel. Further, the first and the fourth terms will also cancel. In particular, the first term solves as

$$\frac{\partial w}{\partial \theta}\int_{a_m}^{a^*} f(a_i)da_i = \frac{\partial w}{\partial \theta}\frac{L^*}{N}$$

For the fourth term, from the envelope theorem we have $\frac{\partial \pi_i^*}{\partial \theta} = -l_i^*\frac{\partial w}{\partial \theta}$. This integral thus solves as

$$\int_{a^*}^{\infty}\frac{\partial \pi_i}{\partial \theta}f(a_i)da_i = -\frac{\partial w}{\partial \theta}\frac{L^*}{N}$$

which cancels with the first integral.

The comparative statics on income per capita thus simplify to

$$\frac{\partial y}{\partial \theta} = w \int_{a_m}^{a^*} \frac{\partial f(a_i)}{\partial \theta} da_i + \int_{a^*}^{\infty} \pi_i \frac{\partial f(a_i)}{\partial \theta} da_i$$

Now, consider the case of a right shift in the distribution $f(a_i)$, where $\frac{\partial f(a_i)}{\partial \theta} > 0$ for all $a_i \geq a^*$. Noting that $\pi_i (a^*) = w$ and $\pi_i > w$ for all $a_i > a^*$, it follows that,

$$\int_{a^*}^{\infty} \pi_i \frac{\partial f(a_i)}{\partial \theta} da_i > \int_{a^*}^{\infty} w \frac{\partial f(a_i)}{\partial \theta} da_i$$

and therefore

$$\frac{\partial y}{\partial \theta} > w \int_{a_m}^{a^*} \frac{\partial f(a_i)}{\partial \theta} da_i + w \int_{a^*}^{\infty} \frac{\partial f(a_i)}{\partial \theta} da_i = w \int_{a_m}^{\infty} \frac{\partial f(a_i)}{\partial \theta} da_i = 0$$

Thus income per-capita is increasing with a right shift in the distribution of entrepreneurial acumen, as was to be shown. Similar reasoning gives the other two cases.

Comparative Statics on Π^*/N

The equilibrium profit rate is such that $\Pi/Y = 1 - \beta$. Thus comparative statics for profits per capita follow the direction as the comparative statics for income per capita, which are shown above.

Proof of Proposition 2

Consider first the equilibrium a^*.

For a Pareto distribution, the share of founders is

$$\frac{E^*}{N} = \int_{a^*}^{\infty} f(a_i) da_i = \left(\frac{a_m}{a^*}\right)^{\gamma},$$

the share of workers is

$$\frac{L^*}{N} = \int_{a^*}^{\infty} l_i^* f(a_i) di = \left(\frac{a_m}{a^*}\right)^{\gamma} \left(\frac{\beta}{w}\right)^{\frac{1}{1-\beta}} \frac{\gamma(1-\beta)}{\gamma(1-\beta)-1} (a^*)^{\frac{1}{1-\beta}},$$

and using the resource constraint, $E + L = N$, we then have

$$1 = \left(\frac{a_m}{a^*}\right)^{\gamma} \left[1 + \left(\frac{\beta}{w}\right)^{\frac{1}{1-\beta}} \frac{\gamma(1-\beta)}{\gamma(1-\beta)-1} (a^*)^{\frac{1}{1-\beta}}\right]$$

Using the entrepreneurial entry condition, (4), to replace w in this expression, we have (after some algebra)

$$a^* = a_m \left[\frac{\gamma - 1}{\gamma(1-\beta) - 1}\right]^{1/\gamma}$$

which is the first result in the proposition, as was to be shown.

The equilibrium wage, w^*, then follows using a^* in (4), giving

$$w^* = a_m \left[\frac{\gamma - 1}{\gamma(1-\beta) - 1}\right]^{1/\gamma} \beta^{\beta} (1-\beta)^{1-\beta}$$

The equilibrium entry rate into entrepreneurship, E^*/N, follows using a^* in (5), giving

$$\frac{E^*}{N} = \left(1 - \frac{\gamma\beta}{\gamma - 1}\right)$$

The equilibrium share of workers is therefore

$$\frac{L^*}{N} = \frac{\gamma \beta}{\gamma - 1}$$

Equilibrium GDP then follows, after some algebra, noting that

$$Y^* = (w^* L^*) / \beta = a_m \left[\frac{\gamma - 1}{\gamma (1 - \beta) - 1} \right]^{1/\gamma} \beta^\beta (1 - \beta)^{1-\beta} \frac{\gamma}{\gamma - 1},$$

completing the proof.

Proof of Corollary 1

Now consider the equilibrium rate of entrepreneurship within a given group. We have

$$\frac{E_j^*}{N_j} = \int_{a^*}^{\infty} f_j (a) \, da.$$

For the Pareto distributions, (3), this integrates as

$$\frac{E_j^*}{N_j} = \int_{a^*}^{\infty} \frac{\gamma a_j^\gamma}{a^{\gamma+1}} da = \left(\frac{a_j}{a^*} \right)^\gamma.$$

Given the equilibrium value of a^* this becomes

$$\frac{E_j^*}{N_j} = \left(\frac{a_j}{a_m} \right)^\gamma \left[1 - \frac{\gamma \beta}{\gamma - 1} \right]$$

as was to be shown.

Proof of Corollary 2

The firm size distribution within a given group follows from $f_j (a)$ and the relationship between the founder acumen and the resulting firm size. The monotonic mapping between employment size and acumen is, where $a_i \geq a^*$,

$$l_i^* = \left(\frac{\beta a_i}{w} \right)^{\frac{1}{1-\beta}} = \frac{\beta}{1 - \beta} \left(\frac{a_i}{a^*} \right)^{\frac{1}{1-\beta}} \tag{15}$$

and the minimum firm size is therefore $l_i^* = \frac{\beta}{1-\beta}$, as was to be shown.

Let the firm size distribution for group j be $g_j (l_i^*)$ and note that the total count of firms from group j is E_j^*. The fraction of firms of a given size, per member of the group's overall population size is then

$$\frac{n_j (l_i^*)}{N_j} = \frac{E_j^*}{N_j} g_j (l_i^*)$$

The firm size distribution, $g_j (l_i^*)$, using the change-in-variables rule, is

$$g_j (l_i^*) = \left| \frac{da_i (l_i^*)}{dl_i^*} \right| f_j (a_i (l_i^*) | a_i \geq a^*)$$

where $f_j (a_i (l_i^*) | a_i \geq a^*) = f_j (a_i (l_i^*)) / (E_j^* / N_j)$. We can then write

$$\frac{n_j (l_i^*)}{N_j} = \left| \frac{da_i (l_i^*)}{dl_i^*} \right| f_j (a_i (l_i^*)) \text{ where } l_i^* \geq \frac{\beta}{1 - \beta}.$$

Inverting the monotonic relationship (15), we have

$$a_i (l_i^*) = a^* \left(\frac{1 - \beta}{\beta} \right)^{1-\beta} (l_i^*)^{1-\beta}.$$

The slope of acumen with firm size is then

$$\frac{da_i\left(l_i^*\right)}{dl_i^*} = (1-\beta)\,\frac{a_i\left(l_i^*\right)}{l_i^*}.$$

And we can thus write

$$\frac{n_j\left(l_i^*\right)}{N_j} = \left(\frac{a_j}{a_m}\right)^\gamma \left(1 - \frac{\gamma\beta}{\gamma-1}\right)\gamma\left(1-\beta\right)\left(\frac{1-\beta}{\beta}\right)^{-\gamma(1-\beta)}\left(l_i^*\right)^{-\gamma(1-\beta)-1}$$

Taking logs produces the result in the corollary, as was to be shown.

Proof of Corollary 3

Define the total number of jobs created by a given group j as M_j. This count is the total number of founders from that group, E_j^*, plus the total number of wage workers in the firms these founders create, which we define as L^{j*}. We are interested in whether M_j exceeds the population size of the group, N_j. We have

$$\begin{aligned}
M_j &= E_j^* + L^{j*} = N_j \int_{a^*}^{\infty} f_j\left(a_i\right)da_i + N_j \int_{a^*}^{\infty} l_i^*\left(a_i\right)f_j\left(a_i\right)da_i. \\
&= N_j \int_{a^*}^{\infty} \left[1 + l_i^*\left(a_i\right)\right]f_j\left(a_i\right)da_i
\end{aligned}$$

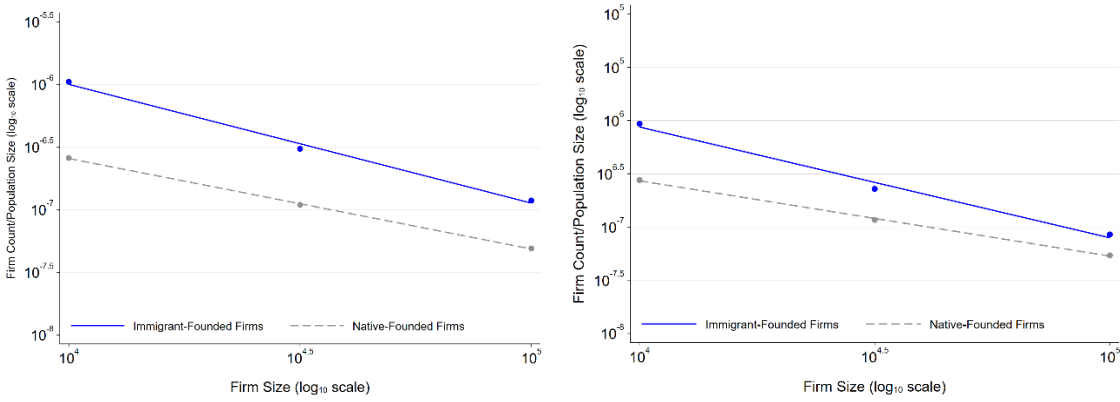Performing the integral for the Pareto distribution equilibrium we find

$$M_j = N_j\left(\frac{a_j}{a_m}\right)^\gamma = N_j\left(\frac{a_j^\gamma}{a_0^\gamma n_0 + a_1^\gamma n_1}\right)$$

Thus the number of jobs created by the group exceeds the group's size iff the Pareto threshold parameter for that group is larger than the for the other group. In particular, for immigrants,

$$\begin{aligned}
M_1 &> N_1 \quad \text{if } a_1 > a_0 \\
M_1 &= N_1 \quad \text{if } a_1 = a_0 \\
M_1 &< N_1 \quad \text{if } a_1 < a_0
\end{aligned}$$

as was to be shown.

Figure B1
Immigrant and Native-Born Entrepreneurship:
Firm Size Distributions using the Fortune 500 firms founded post-1970

Notes: Each panel consider the firm size distribution, distinguishing between immigrant-founded and native-founded firms. The x-axis is the log of firm size measured as current total employment in the firm, using the 2017 Fortune 500. The y-axis is the log count of firms of a given size, with the count normalized by the population size for the relevant group (immigrant or native born). The population measure is an average of the immigrant or native-born population in the decade of founding, weighted by the number of firms founded in that decade. The plotted measures correspond to equation (12). Panel A counts a firm as immigrant-founded if any of the founders are immigrants (Definition 1 in the text). Panel B assigns firms to immigrant and non-immigrant proportionally based on the mix of immigrant and native-born founders of the initial business (Definition 3). Definition 2 is not available for the Fortune 500, as discussed in text.

Figure B2
Technology-Based Immigrant and Native-Born Entrepreneurship
using Definition 2
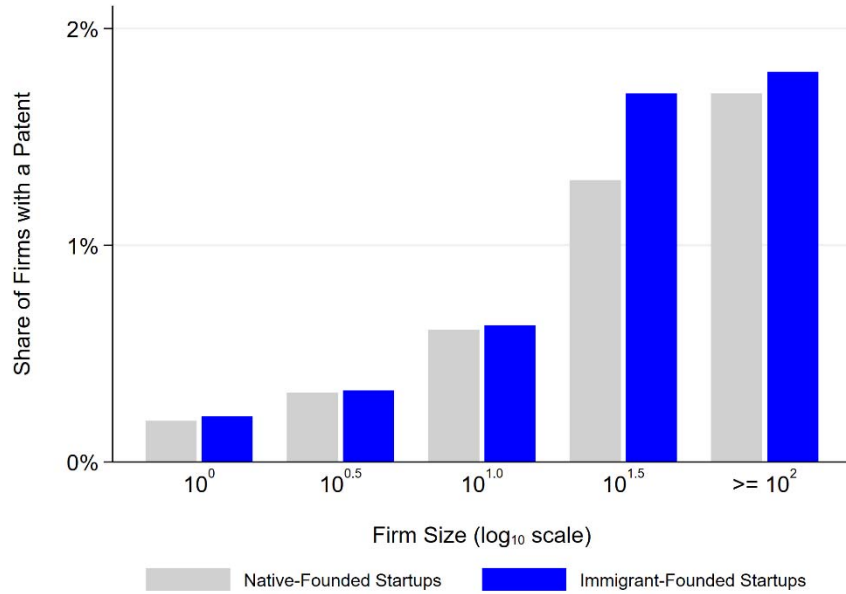
Notes: Using W-2 and LBD data combined with patenting records from the USPTO, this figure shows the share of firms in each firm size bin that own at least one patent, distinguishing between native-founded versus immigrant-founded startups, for all firms in the US between 2005 and 2010. Immigrant-founded startups are identified using Definition 2, which equals 1 if the highest paid founder is foreign-born. Firms are grouped into five bins along the x-axis based on the number of employees five years after founding. The difference in firm size binning relative to Figure 4 is due to Census disclosure rules requiring the minimum number of firms represented in each cell.

Table B1: Wages at Immigrant versus Native-Founded Firms

	DV = Ln(Annual Wages)					
	(1)	(2)	(3)	(4)	(5)	(6)
Immigrant-founded firm	0.041***	0.045***	-0.000	-0.040***	0.008***	0.007***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Ln(Firm size)						-0.029***
						(0.000)
Male				0.337***	0.245***	0.246***
				(0.001)	(0.001)	(0.001)
Foreign born				0.082***	0.124***	0.121***
				(0.001)	(0.001)	(0.001)
Observations (Individuals)	14,640,000	14,640,000	14,640,000	14,640,000	14,640,000	14,640,000
R^2	0.000	0.001	0.010	0.131	0.230	0.231
Individual Age Fixed Effects				YES	YES	YES
Founding Year Fixed Effects		YES	YES	YES	YES	YES
County Fixed Effects			YES	YES	YES	YES
NAICS-4 Fixed Effects					YES	YES

Notes: This table shows a series of OLS regressions using log annual wages as the dependent variable. Sample consists of individuals employed by startups at five years after founding, distinguishing immigrant versus native-founded firms based on Definition 1. Standard errors in parentheses are clustered at the firm level. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$

Appendix References

Gibson, Campbell, and Emily Lennon. 1999. "Historical Census Statistics on the Foreign-Born Population of the United States: 1850 to 1990." *Working Paper No. 29*, U.S. Bureau of the Census, Population Division.

Passel, Jeffrey S., and D'Vera Cohn. 2018. "U.S. Unauthorized Immigrant Total Dips to Lowest Level in a Decade." Pew Research Center. Available at `https://www.pewresearch.org/ hispanic/2018/ 11/27/u-s-unauthorized-immigrant-total-dips-to-lowest-level-in-a-decade/`.

Van Hook, Jennifer, and James D. Bachmeier. 2013. "How Well Does the American Community Survey Count Naturalized Citizens?" *Demographic Research* **29**(1): 1-32.