I Z A Institute of Labor Economics

Initiated by Deutsche Post Foundation

DISCUSSION PAPER SERIES

IZA DP No. 11662

High-Growth Entrepreneurship

J. David Brown
John S. Earle
Mee Jung Kim
Kyung Min Lee

JULY 2018

DISCUSSION PAPER SERIES

IZA DP No. 11662

High-Growth Entrepreneurship

J. David Brown
U.S. Census Bureau and IZA

John S. Earle
George Mason University and IZA

Mee Jung Kim
U.S. Census Bureau and George Mason University

Kyung Min Lee
George Mason University

JULY 2018

ABSTRACT

High-Growth Entrepreneurship*

Analyzing data on all U.S. employers in a cohort of entering firms, we document a highly skewed size distribution, such that the largest 5% account for over half of cohort employment at firm birth and more than two-thirds at firm age 7. Little of the size variation is accounted for by industry or amount of finance, but relative size is strongly persistent over time: at age 7, the probability of 20+ employees is about 40 times larger for those entering with 20+ than for those entering with one. We link administrative and survey data to study the role of founder characteristics in high growth, defined as the largest 5% of the cohort at ages 0 and 7. Female-founded firms are 50% less likely to be in this ventile at both ages, and 34% less likely when controlling for detailed demographic and human capital variables. A similar initial gap for African-Americans, however, disappears by age 7. Founder age is positively associated with high growth at entry, but the profile flattens and turns negative as the firm ages. The education profile is initially concave, with graduate degree recipients no more likely than high school graduates to found high growth firms, but the former nearly catch up to those with bachelor's degrees by firm age 7, while the latter do not. Most other relationships of high growth with founder characteristics are highly persistent over time. Prior business ownership is strongly positively associated, and veteran experience negatively associated, with high growth. A larger founding team raises the probability of high growth, while, controlling for team size, diversity (by gender, age, race/ethnicity, or nativity) either lowers the probability or has little effect. Controlling for start-up capital raises the high-growth probability of firms founded by women, minorities, immigrants, veterans, smaller founding teams, and novice, younger, and less educated entrepreneurs. Perhaps surprisingly, female, minority, and less-educated entrepreneurs tend to choose high-growth industries, but fewer of them achieve high growth relative to their industry peers.

JEL Classification: D22, J24, L25, L26

Keywords: entrepreneurship, business entry, firm growth, firm dynamics, founder, employment, firm size distribution, firm performance

Corresponding author:
John S. Earle
School of Policy
George Mason University
3351 Fairfax Drive, MS 3B1
Arlington, VA 22201
United States
E-mail: earle@gmu.edu

* We benefitted from comments of seminar participants at the Census Bureau, Federal Reserve Board, National Bank of Hungary, Paris-Dauphine, ZEW, IZA, and SOLE, and from Zoltan Acs, Ryan Decker, Nathan Goldschlag, Javier Miranda, and Martha Stinson. We thank the National Science Foundation for support (Grants 1262269 and 1719201 to George Mason University). Any opinions and conclusions expressed herein are ours only and do not necessarily reflect the views of the NSF or Census Bureau. All results have been reviewed to ensure that no confidential information on individual firms is disclosed.

1. Introduction

Recent research exploiting the availability of large firm-level datasets has made great strides in understanding employment patterns by firm size and age in the U.S. Conclusions about the role of small versus large firms dating back to Birch (1979, 1981, 1987) have been amended to recognize the predominance of entrants and young firms in net job creation (Haltiwanger, Jarmin, and Miranda 2013; Decker et al. 2014). At the same time, there has been increasing recognition that most firms enter at a small size and remain small afterward (Shane 2008, Hurst and Pugsley 2011, 2017). While these empirical regularities may seem mutually inconsistent, they can be reconciled if firm size and growth positively skewed, so that a small fraction of all entrants account for most job creation in a cohort.[1]

The importance of this high-growth entrepreneurship is widely recognized, yet many open questions remain. In this paper, we study two sets of such questions. The first set concerns the basic patterns of size at entry and subsequent growth. What fraction of employment is accounted for by the right tail of high-growth entrants? Do such firms begin operations already at an unusually large size, or are they initially indistinguishable from other entrants and only become large after several years of rapid growth? To what extent is heterogeneity in start-up size and growth accounted for by the industries in which firms operate, and to what extent by the availability of finance? Much of the literature has focused either on finance or on differences across industries, but even after controlling for narrow industries and finance we find substantial size variation both at start-up and subsequently.

This heterogeneity motivates a second set of questions about the characteristics of the owner-founders of high-growth entrants. Do the founders of these firms differ from others by demographic characteristics such as age, gender, race/ethnicity, and citizen/immigrant status? Does the human capital of high-growth entrepreneurs differ in terms of education, general labor market experience, veteran status, and prior entrepreneurial experience, compared to owners of low-growth firms? Are the founding teams of high-growth entrepreneurships larger, and to what extent do they involve family members versus unrelated individuals or with more diverse founding teams, defined by age, gender, and race/ethnicity? Do these patterns vary if the amount of start-up finance and specific industry choice are taken into account? Finally, how persistent are the impacts of start-up characteristics on the probability of high growth as the firm ages?

We address these questions following a cohort of firms from their initial entry and analyzing a large, representative data set containing a rich set of characteristics. Rather than study incumbent firms that have already attained some minimum size, as is common in research on "high-growth firms," we instead focus on high-growth *entrepreneurship* by tracking firms from the first quarter in which they hire an employee and analyzing the determinants of being in the top five percent in employment size at age zero (their entry quarter) and at age seven (28 quarters later). Our analysis in each case thus compares firms at exactly the same age, focusing on start-up through age seven. We avoid conditioning on prior growth, and at age 7 we treat early growth and

[1] Much of what we know about the firm size distribution comes from studies of cross-sectional data on existing firms; studies of entrants include Cabral and Mata (2003) on Portuguese manufacturers; Lotti and Santarelli (2004) on four manufacturing industries in Italy; and Angelini and Generale (2008), also on Italy (although they pool all firms up to age six). Section 2 below provides further information on these papers and other related research.

later "catch-up" symmetrically: all jobs created by firms from their initial entry are counted, rather than excluding those created at start-up or through some later age.

The data we analyze include the Business Register (BR) and Longitudinal Business Database (LBD), covering the universe of U.S. private employers, for analyzing the patterns of entry and growth. In order to incorporate founder characteristics we focus on the 2007 entry cohort and link to the 2007 Survey of Business Owners (SBO), resulting in about 37,100 observations on start-up firms and 55,800 on founders. The rich set of founder characteristics in these data goes beyond the basics of age and years of schooling considered in studies such as Cabral and Mata (2003). In addition, we analyze the roles of gender, detailed race/ethnicity, type of schooling, and other aspects of human capital: veteran status, citizen/immigrant, and previous entrepreneurship. Exploiting detailed information on up to four owners of each firm, we study the size and composition of founding teams of entrepreneurs, including the extent to which diversity is correlated with high growth. Linking to the BR permits us to track this entry cohort until age 7, the last available year. To check whether results are sensitive to the macroeconomic environment, we also carry out an analysis of the 2012 entry cohort, as one of many robustness checks.

The remainder of the paper is organized as follows. The next section explains how our paper relates to previous research. Section 3 describes our data and measurement approach. Section 4 contains results, first describing the empirical regularities of size at entry and the transitions between size categories with age, and then providing estimates of the impact of the founder characteristics, start-up finance, and narrow industry on the probability of a firm being in the top ventile of the employment size distribution. The concluding section summarizes and draws out some further implications of the findings. Appendices include detailed data description (Appendix A), full regression results for estimates shown graphically in the text (Appendix B), and robustness checks (Appendix C).

2. Previous Research

No previous research addresses quite the over-arching question of this paper: what sort of entrepreneurs are most likely to found firms at the far right tail of the job creation distribution? This paper therefore builds on several strands of previous research, but it also differs in fundamental ways. Our question focuses not on incumbent firms and their owners but on start-ups and their founders, and on the right tail of job creation rather than the mean of the distribution. Our data set is not only large and representative but also permits us to examine and control for a much larger set of interesting founder characteristics than previous research (mostly on mean effects for incumbents) has been able to analyze. In this section, we explain how our approach relates to existing knowledge.

Our starting point is Haltiwanger, Jarmin, and Miranda's (2013) finding that the only age group with substantial positive net job creation is entrants.[2] We add to this finding in several ways. Unlike their examination of mean differences by age and size, we focus on the top ventile, we consider the same set of firms at different ages and compare job creation across them, and we investigate whether large entrants tend to remain large and the extent to which they continue to

[2] Employment among entrants in their 2005 data is 3.5 million, all of which is job creation by definition. The only other age group with positive net job creation is over 25 years old. They report net job creation for this latter group at 400,000, which can be compared to a total employment of 6.9 million that same year (U.S. Census Bureau, 2016).

grow. In addition to looking beyond the mean and in tracking an entry cohort, we also analyze the association of high job creation with founder characteristics.

The research in this paper is also connected to empirical studies of the firm size distribution. Most of that literature examines cross-sections of existing firms, rather than entrants (Sutton 1997 contains a review). Prominent exceptions include Cabral and Mata (2003), who study the evolution of the size distribution in Portuguese manufacturers; Lotti and Santarelli (2004), who analyze four manufacturing industries in Italy; and Angelini and Generale (2008), who study the size distribution by age in Italy, although they pool all firms up to age six. Sample sizes tend to be small in these studies, and their focus is generally on estimating how the overall distribution evolves with firm age. They do not examine the right tail, including the fraction of employment accounted for by the largest firms, and they do not track the relative size of individual firms and the changes in their position in the size distribution. Except for Cabral and Mata (2003), who examine the relationship between size and owner age and education, these studies do not examine characteristics of founders, and they do not relate the characteristics to the probability of job creation on the right tail of the distribution.

Our focus is on high growth during the initial, entrepreneurial phases of entrants, but the analysis is related to literature on "high-growth firms." Most of this research is essentially cross-sectional in comparing firms without regard to age or stage of life cycle. Most of it ignores start-ups entirely and examines only existing firms, incumbents. Some researchers define "high growth" as the top 1, 5, or 10 percent of the growth rate distribution, but this tends to result in a bias towards initially small firms, while defining growth in absolute terms biases toward large firms. Some focus on growth in a particular year, although as noted below, growth is highly volatile over time so that a particular year may not reflect longer term job creation.[3] To address these problems, the Eurostat and OECD (Eurostat-OECD 2008; OECD 2010) propose a definition of high-growth firms as those with at least 10 employees at a certain time and an average of at least 20 percent annual growth over the next three years. As Daunfeldt, Johansson, and Halvarsson (2015) point out, however, the 10-employee initial size restriction excludes the vast majority of firms.[4]

These approaches to defining "high growth firms" as incumbents with a high growth spurt over some time period also face the problem that firm growth is extremely volatile even with respect to multi-year periods (e.g., Acs, Parsons, and Tracy 2008; Holzl 2014; McKelvie and Wiklund 2010). Daunfeldt and Halvarsson (2015), for example, show that Swedish firms with a three-year period of high-growth tend to have declining growth in the previous three-year period, and the probability that they repeat their high growth performance in either of the next two three-year periods is very low. These high growth definitions also exclude job creation from entrants.

Other studies focus on start-up size, and a subset of those follow the same cohort of firms for several years from start-up.[5] Coad et al. (2014) suggest that the best way to ensure a firm

[3] Decker et al. (2014) defines high-growth for incumbents as employment increase over 25 percent, Stangler (2010) examines the top one and five percent, and Storey (1994) the top four percent without regard to age.

[4] They report that the initial size restriction excludes almost 95 percent of surviving firms in their Swedish sample. Another approach is Acs, Parsons, and Tracy (2008) definition of "high-impact firms" as those at least doubling sales over a four-year period and with a product of absolute and percent change in employment (sometimes called the "Birch index") of at least two during the same period.

[5] In addition to the size distribution studies cited above, see Garnsey et al. (2006), who argue that following a cohort reduces survival bias and may increase consistency in the measured impacts of firm growth factors.

reaches a large size at a particular age is to be large at start-up. As we show below, start-up size is a powerful indicator of size at age 7. Examining employment at age 7 size places uniform weight on job creation throughout the entrepreneurial period, including from start-up. By this measure, firms can be high growth either by creating many jobs at start-up or by catching up later.[6]

Concerning the determinants of high growth, a long literature in management and related disciplines has examined some characteristics (e.g., Kalleberg and Leicht, 1991). Again, these are typically cross-sectional analyses of small samples, however, and in many cases they study incumbents and take no account of firm age. Within economics, most studies of firm growth focus on the mean, as in Neumark, Wall, and Zhang (2011) and Haltiwanger, Jarmin, and Miranda (2013).[7] Fairlie and Robb (2007, 2009) and Robb and Fairlie (2009) analyze the impact of race, gender, and family history using the 1992 Characteristics of Business Owners (CBO), the predecessor of the SBO, but they do not observe employment level or growth in these data. Their analysis estimates cross-sectional differences in mean sales, survival, and the probability of hiring at least one employee, without distinguishing by firm age, all of which differ from our focus on high-growth entrepreneurship among entrants.[8]

Previous research studying owner characteristics with a cohort of entrants includes Cabral and Mata's (2003) analysis of Portuguese firms and studies of the Kauffman Firm Survey (KFS) in the U.S. Cabral and Mata (2003) study entrants at age 0 and 7, as do we, but they do not focus on the high-growth group. In analyzing characteristics, they condition on survival to age 7, so that firms exiting before age 7 are not in their age 0 analysis. Their sample is restricted to manufacturing, the sample size is 515 firms, and as discussed above the owner characteristics are only age and education. The KFS is a cohort and it includes rich information on founders, but the samples are very different in size and composition from ours: the KFS sample contains fewer than 5,000 entrants, and the sample is drawn from a list of Dun and Bradstreet firms, which is more likely to include firms that already had some credit history, unlike our data, where inclusion is based on reporting payroll employment to the Internal Revenue Service. The KFS sample includes nonemployers and purchases of existing businesses and purchases of franchises, which we exclude.[9] Fairlie and Miranda (2017) use these data to examine the determinants of a non-employer becoming an employer. However, it would be difficult to do the same with the KFS data for an analysis of the probability of being in the right tail of job creation because the sample in that tail would be so small, preventing reliable estimation of the association with founder characteristics, a general problem for research on high growth in entry cohorts.

Similar to our research in terms of data are Jarmin and Krizan (2010) and Jarmin, Krizan, and Luque (2014), who link data on firm characteristics from the 2002 SBO to a longitudinal data

[6] Our approach stands in contrast to the bulk of studies that count only post-entry growth. For instance, in a study of mean employment growth in immigrant-owned firms, Kerr and Kerr (2017) also follow cohorts of entrants, but they exclude age 0 job creation from their measure of employment growth (and their regression estimates control for age 0 employment).

[7] Also related are studies of heterogeneity in firm performance, such as Bloom and van Reenan (2007) and Syverson (2011), but again the focus is on incumbent firms, generally large corporations, with little or no attention to entrants and founder characteristics.

[8] Bates (1990) examines firm survival using the 1982 CBO data.

[9] In another type of study, Garnsey et al. (2006) examine growth patterns in cohorts including about 400 firms, and Brown et al. (2005) examine growth determinants in a similarly sized sample.

source on employment (LBD).[10] They focus on mean, not high, growth determinants, and they analyze the cross-section of all firms, rather than an entry cohort. The data do not permit them to study several important issues including immigrant status, husband-wife ownership, prior business ownership, and amount of start-up capital, which we are able to address with the 2007 SBO.[11]

Despite these substantial differences between our approach and the previous research, we discuss some of the key results from this literature to provide context when we report our findings below. Similar to the previous research, our aim is to establish important empirical regularities that may be useful for theory and policy, but not to test an explicit model. The results are relevant for some theories, however. Our finding of large heterogeneity in firm size at entry, even within narrowly defined industries, is at odds with canonical models of industry dynamics going back to Jovanovic (1982) and Hopenhayn (1992) and extending to Melitz (2003) etc., which have all entrants choosing the same optimal size (Frank 1988 is an exception). Our result that entrant size heterogeneity declines when start-up capital is taken into account is suggestive that varying financial constraints may account for some of the size heterogeneity. On the other hand, our finding of high persistence of size from age 0 to age 7 suggests some strong underlying heterogeneity in firm potential and possibly in founder motivations that deserves further research. We find the "up-or-out" dynamic, a productivity-enhancing mechanism frequently posited in discussions of industry dynamics (e.g., Decker et al. 2014), is strongest among the largest entrants, and is weaker in smaller start-ups, which tend to remain about the same size.

Our results also relate to theoretical concepts of human capital, discrimination, occupational choice, and complementarities. The finding that more education is not uniformly valuable in raising firm success (defined by employment size) challenges single-factor models of human capital, and suggests instead that multiple dimensions of skill are relevant. The finding of lower prevalence of women and minorities among high-growth entrepreneurs could be consistent with theories of either discrimination or self-selection into occupations, but the result that these differences are diminished when start-up capital is taken into account suggests possible discrimination in financial markets. A further result that the gaps are larger within narrow industries implies, contrary to the possibility that women and minorities choose unambitious fields in which to open businesses, that in fact they choose high-growth sectors, but their performance is worse (in the sense of firm size) within sector.

Finally, the research in this paper is relevant to notions of complementarity and diversity within teams. Are larger teams more likely to found high-growth firms (Ruef, Aldrich, and Carter 2003)? What kinds of skills and characteristics combine to promote growth? Lazear (2005) has posited the desirability of "balanced skills" for individual entrepreneurs, but perhaps the balance can be achieved with a diverse team. Again, these issues have previously been studied at the mean for a cross-section of incumbents, while our focus is on the right tail of growth among start-ups.

---

[10] Fairlie and Robb (2007, 2009) and Robb and Fairlie (2009) use the CBO to study determinants of a firm being an employer at the time of the survey. By contrast, Jarmin and Krizan (2010), Jarmin, Krizan, and Luque (2014), and our paper use data on the number of employees.

[11] In an important contribution appearing after we drafted this paper, Azoulay et al. (2018) focus on founder age as a determinant of high growth at firm age 5. They do not examine or control for other demographic, human capital, and team characteristics or for the amount of start-up capital, as we do in this paper, but they are able to look at a variety of success measures, including successful exits through IPO, which are not observed in our data.

3. Data and Methods

3.1 Data

This subsection contains a brief description of the data. More details as well as descriptive statistics for the variables are provided in Appendix A.

The basic sources for tracking employment are the Census Bureau's Business Register (BR) and Longitudinal Business Database (LBD). The quarterly BR includes all nonfarm businesses filing Internal Revenue Service tax forms as individual proprietorship, partnership, or corporation, and with receipts of $1,000 or more. We use the longitudinal establishment links from the LBD to track firms and their reorganizations over time. We aggregate establishments to obtain firm-level employment and define age = 0 as the firm's first quarter with positive employment.

We focus on the four quarterly entry cohorts in 2007 and link these data to the 2007 Survey of Business Owners (SBO), which is a large random sample from the BR. The SBO is quinquennial, and we use the data from 2007 rather than 2002 or 2012 because the 2002 SBO lacks information on several of the factors we wish to study, and the 2012 SBO permits observation only on a short time span after start-up (2014 is the most recent available year in the BR). An obvious concern is that results may differ for firms founded in 2007, just before the Great Recession, compared to firms founded in other years, so we have also estimated all the Age 0 relationships with the 2012 data. The results, discussed in the robustness sub-section and provided in Appendix C, are very similar to those from 2007.

We are interested in studying the determinants of job creation over a longer period, not only at start-up. For this purpose, we use age 7 employment, defined as the firm's employment in the same quarter of 2014 as its start-up quarter in 2007. We choose age 7 for measuring the firm's longer-run net job creation because it is the oldest age we can currently observe for the 2007 start-up cohorts (2014 is the most recent available year for the BR), and because some researchers (e.g., Nightingale and Coad 2014) define "entrepreneurial firms" as those under age 7. Thus, age 7 employment is the net job creation over the entrepreneurial period so defined. Even if available, using employment at a later age would have drawbacks: the older the firm, the more difficult it is to attribute its performance to a single origin and founding team, because of firm boundary and ownership changes taking place over time.

Appendix A describes in detail our procedures for tracking firms longitudinally, taking into account changes in identifiers. The problems are less acute in our cohort of entrants than among older and larger firms.[12] The Appendix also describes our measurement of organic growth, removing the effects of establishment acquisitions and divestments.

The linked BR-SBO sample consists of 55,800 owners of 37,100 firms, about 7.0 percent of all firm start-ups that year (U.S. Census Bureau, 2016). We weight each owner by ownership equity shares (summing to one for each firm). We weight each firm by the inverse of the sample-population ratio (the share of firms in the two-digit NAICS industry–employment category in the LBD divided by the sample's share of firms in the two-digit NAICS industry–employment category).

All the independent variables are measured for the year 2007, the start-up year for the firms in the main sample and the reference year in the SBO. Founder characteristics from the SBO

[12] McKelvie and Wiklund (2010) discuss challenges of tracking firms across time when measuring growth.

include basic demographics (age, gender, race, ethnicity, and immigrant/native), human capital (type of education, veteran, and prior entrepreneurial experience), and the size of the founding team and relationships among multiple founders (family/unrelated and diversity by demographics and education). We use the firm's 6-digit industry from the BR, and categories of the amount of start-up finance from the SBO (0-5, 5-10, 10-25, 25-50, 50-100, 100-250, 250-1mln, >1mln, in 1000s of $; a small number respondents answered "unknown" or "no capital needed," which we control for but do not report).

Means of the finance variables and founder characteristics are provided in Appendix Table A1, as are details of the construction of characteristics measures from the raw data. An important point is that for firms with multiple owners, the gender, race, ethnicity, and immigrant variables are defined to indicate whether all the firm's owners are in that category, and thus include a label "all," in order to permit us to measure the impact of diversity. In the case of gender diversity, for instance, we define "all female" and "all male" variables to indicate firm with owners from only one gender or the other (including single owners). The gender diversity variable indicates that the business is jointly owned by at least one owner of each gender (except when husband and wife, for which we provide a separate category). By controlling for all-female, the variables for different types of husband-wife ownership and gender diversity for non-couples measure whether gender ownership effects vary depending on who else co-owns the firm. Racial/ethnic diversity indicates that the business is jointly owned by at least two individuals with different race or ethnicity from one another, immigrant diversity indicates the business is jointly owned by individuals who are immigrants and U.S.-born, and multi-generation indicates that at least one owner is 20 or more years older than another. Similarly, including variables for all of one ethnicity or race, all immigrant, and ethnic/racial and immigrant diversity in the regression allow us to examine whether race/ethnicity and immigrant effects differ with homophily or diversity among founder teams along those dimensions.

3.2 Methods

In this paper we define high-growth entrepreneurship as the subset of entrants with the highest net job creation. Given that entrants have zero employment prior to entry, their net job creation is simply their size. We measure high growth by this definition at ages 0 and 7. In most of the results reported below, the high-growth group is defined as the top ventile (5 percent), distinguishing them from the bottom 95 percent of the employment distribution within the sample. Average employment among those in the top ventile is much larger than in the bottom 95 percent: for the sample studied below, average employment in the top group is 57 and 67, and for the bottom 95 percent it is 3 and 2, at age 0 and at age 7, respectively. The top ventile accounts for 52.0 percent of the sample's employment at age 0 and 66.6 percent at age 7.[13] At age 0 the top ventile has 17 or more employees, and at age 7 the threshold is 19.

The regression specifications are variants of the following equation:

[13] As a robustness check, we have also re-run the analysis using the top 2 percent (36 employees or more at age 0 and 40 employees or more at age 7) and top 10 percent (10 employees or more for both ages) thresholds. The top 2 percent make up 39.4 percent of sample employment at age 0 and 50.8 percent at age 7, while the top 10 percent account for 63.1 percent of employment at age 0 and 79.7 percent at age 7. Results using these alternative thresholds, available upon request, are qualitatively similar to those using 5 percent.

$$\Pr(HG_{ijt}) = X_{ij}\beta + \theta_j\delta + \rho_j + K_j + S_j + u_{ijt}, \tag{1}$$

where HG_{ijt} is a dummy equal to 1 if founder i owns firm j with employment in the top ventile of the employment distribution among firms at age t, and t = 0 or 7 in alternative specifications. X_{ij} contains characteristics for founder i of firm j, θ_j contains firm-level characteristics, ρ_j is a vector of dummies denoting the quarter of 2007 in which firm j first has positive employment, K_j is a vector of start-up capital amount categories, S_j is a vector of 6-digit NAICS industry dummies, and u_{ijt} is an idiosyncratic error term. As noted above, the regressions are weighted by owner shares (so that each firm rather than each owner receives equal weight) and by LBD weights (so that results reflect the full population). We report estimates by linear probability model (LPM), and also logits as a robustness check.

For each firm at each age (0 and 7), we estimate a base specification including only factors that are predetermined at the time of start-up, excluding start-up capital K_j and industry dummies S_j. We exclude these variables from the base specification because they may be at least partly choice variables of the entrepreneur in the start-up process. The firm's growth potential may influence the amount of financing through, for example, the quality of the business plan presented to investors. And entrepreneurs desiring to create high-growth firms may choose sectors where large, fast-growing firms are more common.[14]

Though the factors we examine other than industry and finance are predetermined at start-up, some of them could be jointly determined with the high growth outcome through unobserved channels, including the founders' motivations and the quality of the entrepreneurial idea. For example, it is possible that some human capital investment decisions are driven by the intention to start-up a high-growth business. It is also likely to be easier to recruit additional founding team members when the business idea has greater potential, which could be reflected in a larger coefficient on multiple owners.

To see if start-up capital and sectoral choice are channels through which predetermined characteristics influence high growth, we estimate an additional specification adding K_j and a second specification with both K_j and S_j. If, say, female entrepreneurs systematically receive less financing, and K_j is positively associated with high growth, then including K_j will raise the coefficient on female owner.[15] If a coefficient rises (falls) after controlling for S_j, that suggests that the particular type of entrepreneur systematically selects sectors with a lower (higher) share of high-growth entrepreneurship. Coefficients after controlling for sector also show the performance of particular types of entrepreneurs relative to their competitors, which is relevant for their long-

[14] Hurst and Pugsley (2011) show that nonpecuniary motives are associated with sectors with more small firms.

[15] Systematic differences in the amount of start-up capital could be due to individual choice or external financial constraints. Some types of entrepreneurs may be more reluctant to use their own resources or take on debt than others, their creditworthiness may be systematically different, or investors may discriminate against some types of borrowers. We are unable to test among these alternatives here, although results including the industry dummies provide evidence on the degree to which individual choice is reflected in the sector in which the business operates. Fairlie, Robb, and Robinson (2016) and Coleman and Robb (2014) report that African-American, Hispanic, and female entrepreneurs systematically use less start-up capital. Blanchflower, Levine, and Zimmerman (2003), Blanchard, Zhao, and Yinger (2008), and Fairlie, Robb, and Robinson (2016) provide evidence of discrimination against African-American and Hispanic entrepreneurs in the small business credit market. Coleman and Robb (2014) find that the loan denial rate does not vary significantly by gender, but female entrepreneurs are less likely to apply for credit due to fear their loan application will be denied, even when controlling for measured creditworthiness.

run viability. We test for the statistical significance of such differences in coefficients across specifications by jointly estimating the equations.

We also test for differences in coefficients across age (age 7 versus age 0) by pooling the data for the two ages, allowing all coefficients to vary by age, and testing for equality of the coefficients for the same variable at the two ages. This permits us to assess the degree to which the predictive power of a coefficient for firm size at age 0 persists or diminishes at age 7.

4. Results

4.1 Entry Size and Growth

In order to motivate our focus on a small group of entrants and to set the context for our analysis of founder characteristics, we first examine basic patterns of heterogeneity in the size of firms upon entry and in their subsequent growth. We measure the size distribution of entrants, the concentration of employment among large entrants, and the extent to which firms that enter large remain large and concentrate employment after seven years. The data construction follows the sample procedures described above, containing all firms in the LBD that first have positive employment in the BR in one of the quarters of 2007. We use a transition matrix across size categories for these entrants from start-up in 2007 (age 0) to age 7 in 2014, with categories defined as 1, 2-4, 5-9, 10-19, and 20 or more employees.[16] Age 0 employment must be positive by definition of entrant, but age 7 equals zero if the firm exits.

Tables 1 and 2 display the results, with row percentages in the former and column percentages in the latter. From the "Column Total" in Table 1, it is evident that most firms start very small: 78 percent have initial employment less than 5. Only 4.1 percent of entrants have 20 or more employees, but they account for over half (54.1 percent) of all age 0 employment. Exit

Table 1. Employment Category Transitions from Age 0 to Age 7: Row Percent

		Age 7								
	Emp Size	0	1	2-4	5-9	10-19	20+	Column Total	Age 0 Share	Age 7 Share
Age 0	1	67.3	16.5	10.8	3.4	1.3	0.8	44.7	7.1	15.9
	2-4	61.8	6.6	18.2	8.4	3.2	1.8	33.4	13.9	23.1
	5-9	59.1	2.4	9.5	16.0	8.9	4.1	11.9	12.2	15.0
	10-19	57.3	1.8	4.0	9.2	17.1	10.7	6.1	12.7	13.0
	20+	56.5	0.6	1.6	2.2	7.6	31.4	4.1	54.1	33.0
	Row Total	63.4	10.0	12.3	6.9	4.1	3.3	100.0	100.0	100.0

Note: Each cell represents the percentage of firms in the age 0 size category in the particular row that transition to the age 7 size category in the column. The Age 0 and Age 7 shares are the age 0 size category's percent of employment at age 0 and age 7, respectively.

rates (shown by Age 7 employment = 0) are uniformly high across age 0 size categories, falling only slightly with size. At age 7, only 3.3 percent of the initial start-ups are in the largest size category (Table 1, "Row Total"), but they account for 60.1 percent of all age 7 employment (Table 2, "Age 7 Emp Share"), reflecting a rising concentration of employment as the cohort ages. Of

[16] These are standard size categories used in a variety of contexts. Kerr and Kerr (2017) also provide a transition matrix for start-up cohorts in the LBD, with a focus on the share of immigrants in each cell, but they do not describe the size distribution at age 0 (the marginal distribution).

those starting with 20 or more employees, 31.4 percent remain in that category at age 7, while most of the rest exit. As shown in Table 2, those remaining in the largest category make up 38.2 percent of that category at age 7 despite being just 1.3 percent of the start-ups. Thus, firms starting large have a much higher propensity to be large at age 7 than firms starting smaller: the probability for firms starting with 20 or more employees is 3 times higher than for firms starting with 10-19

Table 2. Employment Category Transitions from Age 0 to 7: Column Percent

		Age 7						
		0	1	2-4	5-9	10-19	20+	Total
Age 0	1	47.4	73.8	39.1	22.2	14.7	10.2	44.7
	2-4	32.5	22.0	49.3	40.8	26.5	17.9	33.4
	5-9	11.1	2.9	9.2	27.6	25.8	14.4	11.9
	10-19	5.5	1.1	2.0	8.1	25.4	19.3	6.1
	20+	3.6	0.2	0.5	1.3	7.6	38.2	4.1
Age 7 Emp Share		0.0	2.8	9.4	12.6	15.2	60.1	100.0

Note: Each cell represents the percentage of firms in the age 7 size category in the particular column that have transitioned from the age 0 size category in the row.

employees, and about 40 times higher than for firms starting with one employee.

Not only do large entrants tend to stay large, but they also tend to grow faster than smaller entrants. Table 3 shows the average employment changes by start-up size category and separately for exiting, declining, unchanging, and growing firms to age 7. The average job loss among exiting

Table 3. Average Jobs Gained/Lost Per Firm by Age 0 Size and Change by Age 7

		Exit	Survive	
		Emp_7=0	ΔEmp<0	ΔEmp>0
Age 0 Size	1	-1.0	N.A.	5.8
	2-4	-2.6	-2.4	8.3
	5-9	-6.5	-5.7	12.0
	10-19	-13.2	-11.2	17.6
	20+	-100.4	-78.5	52.3

Note: Emp7 is employment at age 7 and ΔEmp is changes in employment between age 0 and 7.

and declining firms, or gross job destruction, increases in initial size, which may not be surprising because the larger entrants have more to lose, and as noted their exit rate is not much lower than for smaller firms. But the average absolute employment growth among growing firms is also increasing in start-up size. Gross job creation (per firm) is highest among the largest entrants, and the big future job creators are more likely to be found among the largest entrants. Thus, the "up-or-out" dynamic of fast growth referred to by Decker et al. (2014) is strongest for the largest entrants, further motivating our analysis of this group of firms.[17]

These results demonstrate the importance of understanding the determinants of starting size. Note, however, that, although the probability of growing large (20 or more employees) after starting small (fewer than 20) is much smaller than the probability of remaining large, the former

[17] Decker et al. (2014, p. 10) also note the difference across size groups, reporting that "the average net growth for young firms is substantially higher for firms that are larger than 20 employees. Such patterns highlight that rapid employment growth among young surviving firms is especially present among larger—or at least not micro-sized—young firms."

nevertheless outnumber the latter in an absolute sense: 61.8 percent vs. 38.2 percent of the large category, respectively. The age 7 employment shares of firms by age 0 employment categories are also more evenly distributed than at age 0. It is thus possible that the factors explaining large size at birth and at age 7 could differ significantly. We examine this below.

Some basic models of industry dynamics imply that all entrants should choose the same, optimal size (e.g., Jovanovic 1982; Hopenhayn 1992), but our results contribute to documenting substantial heterogeneity in initial size. To measure the extent to which start-up size and growth can be explained by industry, we estimate a set of regressions of age 0 and age 7 employment on highly disaggregated (6-digit NAICS) industry dummies (measured at age 0). To assess the role of capital in accounting for size variation, we estimate similar regressions with dummies for categories of the amount of start-up capital. We also estimate a set with both industry dummies and start-up capital. The R^2 with industry dummies is 0.086 at age 0 and 0.034 at age 7, it is 0.020 at age 0 and 0.006 at age 7 with start-up capital, and it is 0.100 at age 0 and 0.039 at age 7 with both industry dummies and start-up capital. It is noteworthy that all the R^2s fall sharply with age, implying that within-industry size variation relative to cross-industry variation increases with age.

Given our focus on high-growth entrepreneurship, we also calculate the R^2 from replacing the dependent variable with a dummy for being in the top 5 percent of the employment distribution at the particular age. Using the top 5 percent dummy, the R^2 with industry dummies is 0.128 at age 0 and 0.075 at age 7, it is 0.072 at age 0 and 0.050 at age 7 with start-up capital, and it is 0.172 at age 0 and 0.122 at age 7 with both industry dummies and start-up capital. In all cases, the R^2s fall with age, although less so than for the employment regressions, and they are higher for these top 5 percent regressions than for employment regressions. Thus, the detailed industry and start-up capital variables do better at distinguishing the high-growth group, and their effects are also more persistent for the top 5 percent. But they explain only a small part of the heterogeneity, further motivating our examination of the effects of founder characteristics in the next subsection.

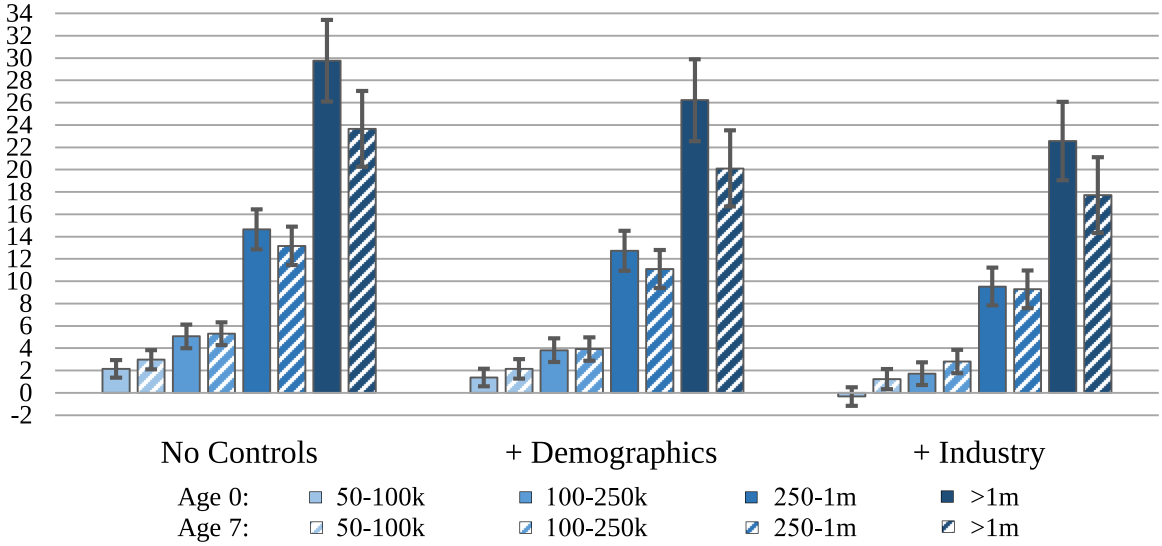
4.2 Founder Characteristics and High Growth Entrepreneurship

For our main analysis of the determinants of high-growth entrepreneurship, we divide firms into the top 5 percent and bottom 95 percent of the sample employment distribution at ages 0 and 7.[18] As a preliminary check before examining the impact of founder characteristics on the probability of being in the 5 percent, high growth group, we first estimate the high growth relationship with the amount of start-up finance and industry in which the firm operates. The start-up finance variable is categorical, as described in Section 2, and the reference category is <$5000. Figure 1 contains coefficient estimates from three specifications of Equation (1): a base specification including finance (K_j) only, a second specification adding demographic, human capital, and team controls (gender, owner age, ethnicity, race, citizenship, education, veteran status, prior business experience, and founding team characteristics – labelled "demographics"), and a third specification also controlling for detailed (6-digit) industry. Results are shown both at firm age 0 (in solid) and age 7 (striped). The Figure does not report results for categories between $5000

[18] As discussed in the data section, the top 5 percent employment thresholds at age 0 and age 7 correspond to 17 and 19 employees, respectively, so results are similar to the 20+ employees category in the transition matrices above. But we find it more natural there to use absolute employment, while here it is simpler to keep the fraction in the top group constant in order to interpret the comparison of results at age 0 and age 7.

and $50,000, which have tiny coefficients, generally statistically insignificant (but full results of coefficients, standard errors, and summary statistics appear in Appendix Table B1). Coefficients

Figure 1. Start-Up Capital and High Growth Probability



Note: LPM regression coefficients for high growth (top ventile). The omitted category for start-up finance is less than 5k, and additional variables not shown are 5-10k, 10-25k, 25-50k, "don't know," and "none-needed." Full results are reported in Table B1. "Demographics" include human capital and team characteristics (Table A1).

and standard errors are multiplied by 100 for ease of reading.

The estimates imply a strong positive relationship between the amount of start-up finance and probability of high growth at both firm ages. For the category of >$1mln, the estimated coefficient in the no controls specification at age 0 implies a 30 percentage point higher probability relative to firms with <$5000, or a probability six times higher than the baseline probability of 5 percent. Controlling for demographics reduces the coefficients only slightly, and controlling for industry a bit more so, but even in the latter case the coefficient of about 23 implies a high-growth probability 4.5 times higher for firms in the largest compared with the smallest finance category. The coefficients rise monotonically in all specifications at both firm ages, but the rise is slower at age 7. The slightly smaller impact of start-up finance after 7 years is inconsistent with initially well-financed firms increasing their financial advantage, but suggests that early and later finance may tend to substitute rather than complement each other in producing high growth. Observed finance may represent a constraint or a choice, but the positive correlation with high growth in these data suggests the measure is useful for our analysis of founder characteristics.

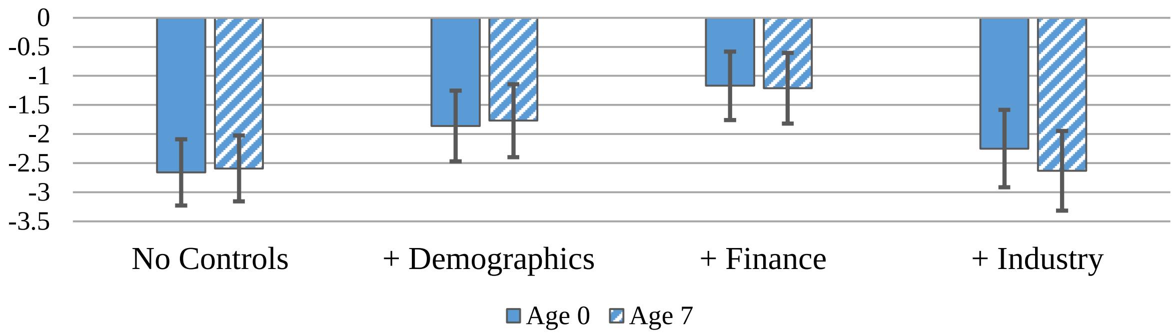
A second firm-level variable we control for in some specifications is 6-digit industry. The probability of being in the top ventile is positive in all sectors, but also varies across them. It is highest in manufacturing. Even more than finance, which is partly determined by founder choices, the industry in which the firm operates is endogenous. We therefore estimate specifications with and without these variables, controlling for them only in order to hold them constant when considering the impact of founder characteristics.

The first characteristic we consider is gender of the founder. Consistent with previous studies that have reported lower rates of business ownership among women, although our methods differ, we find that all female-owned businesses account for a much smaller fraction (16 percent) of employers compared with all male-owned (51.4 percent) in the entry cohort.[19] As discussed above, we distinguish firms with all founders of one characteristics from those with founders of mixed characteristics, and the remaining 32.6 percent of entering businesses have teams of founders containing both genders, as discussed further below.

Estimates of the impact of gender on the probability of high-growth entrepreneurship are shown in Figure 2 (with full estimates in Appendix Table B2), based on Equation (1). Coefficients for all female are negative in all specifications for both firm ages. The magnitude of the estimated effects are not significantly different at age 7 and age 0, and with no controls they imply about a 2.5 percentage point lower probability of being in the top 5 percent high growth group, or a lower probability (relative to the baseline 5 percentage points) of 50 percent. With "demographic" controls, the coefficient falls to 1.7 percentage points, implying a difference of about 34 percent.

This large estimated gender gap falls significantly, to 20 percent, in the third specification shown in the Figure, which controls for amount of start-up finance. This result might be explained by either greater financial constraints for women or more non-pecuniary motives for women

Figure 2. The Gender Gap in High Growth Entrepreneurship



Note: LPM regression estimates also include gender diversity (results in Figure 9). Full results are reported in Table B2. "Demographics" include human capital and team characteristics (Table A1).

founding businesses. Figure 1 showed that the propensity to be a high-growth firm is positively associated with the amount of start-up capital, implying that if a coefficient on an owner characteristic increases (decreases) after controlling for start-up finance, that indicates that the characteristic is associated with less (more) start-up capital. The results here thus indicate that women-owned businesses use less start-up capital, consistent with Fairlie and Robb's (2009) finding that less start-up capital helps to explain lower average sales of women-owned businesses. The magnitude of the age 7 coefficient drops almost as much as the age 0 coefficient when controlling for finance, suggesting that less start-up capital has long-lasting effects on growth.

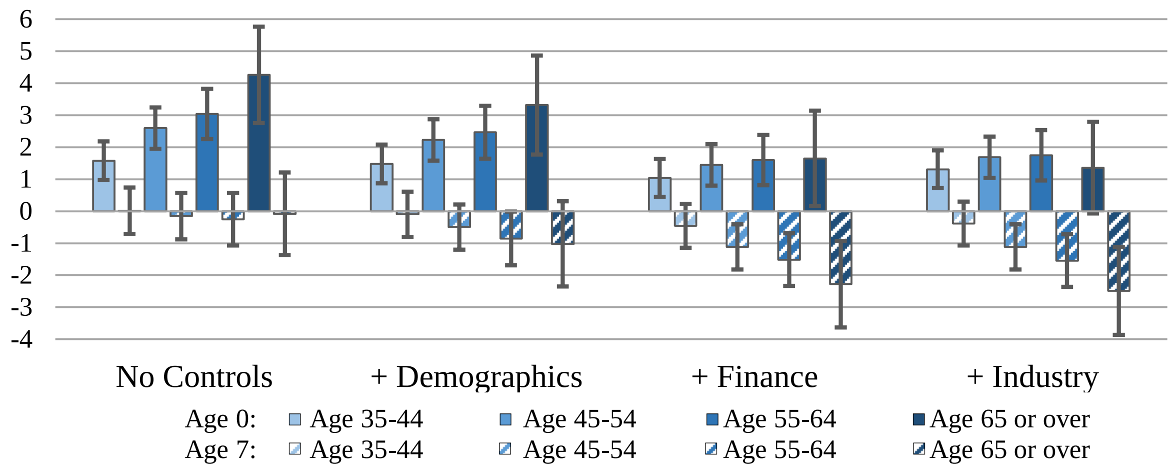
The estimated gender gap increases substantially in the specification controlling for 6-digit industry. The coefficients are similar to those with no controls, implying an approximate 40 percent lower probability of being in the top 5 percent, although they are less precisely estimated.

[19] Jarmin and Krizan (2010) also find that women-owned businesses have lower average employment growth rates in the 2002 SBO linked to the LBD. Using the 1992 CBO, Fairlie and Robb (2009) report women have a lower hiring probability among other measures of business success. Kalleberg and Leicht (1991) find small, statistically insignificant disadvantages of women in survival and earnings growth in a survey of 411 firms, 99 owned by women. These studies examine cross-sections, not distinguishing by firm age, while we follow an entry cohort, and they report average differences, while our focus is the right tail of the distribution.

This result suggests that women tend to choose industries with higher shares of high-growth entrepreneurship, and it appears inconsistent with women starting businesses for non-pecuniary, non-growth-related reasons.[20] It also contrasts with Fairlie and Robb's (2009) report that female businesses have an unfavorable industry distribution for average sales, although again their methods are quite different from ours.[21]

The founder's age may be associated with business success as a result of human capital accumulation (labor market experience, increasing in age), financial constraints (likely to decrease

Figure 3. Founder Age and High Growth Entrepreneurship



Note: LPM estimates. The omitted founder age category is under 35. Full results are reported in Table B2.

with age), and time horizon (decreasing in age). In all cases, the effects may be non-linear. Our results for founder age are presented in Figure 3 (with full estimates in Appendix Table B2). We find that firms with older owners are much more likely to be large at start-up, but the effect disappears by firm age 7.[22] However, we also find that controlling for finance reduces the owner age effect at both firm ages. In this specification at firm age 7, the age profile actually becomes significantly negatively sloped, so that entrepreneurs under 35 years old are much more likely to start top 5 percent companies than those 45 and older. This implies that not only do younger entrepreneurs use less start-up capital, but the negative effect of lower capital is persistent through this early phase of business life. This result is inconsistent with the interpretation that a diminished founder age effect as the firm ages means that liquidity constraints lessen over time.[23] It also suggests that older entrepreneurs may have a skill disadvantage when it comes to starting up high-

[20] Hurst and Pugsley (2011) report a positive correlation between non-pecuniary motivations for founding the business and the share of small firms in the industry.

[21] Fairlie and Robb's (2009) female coefficient changes from -0.69 to -0.57 when controlling for both start-up capital and industry.

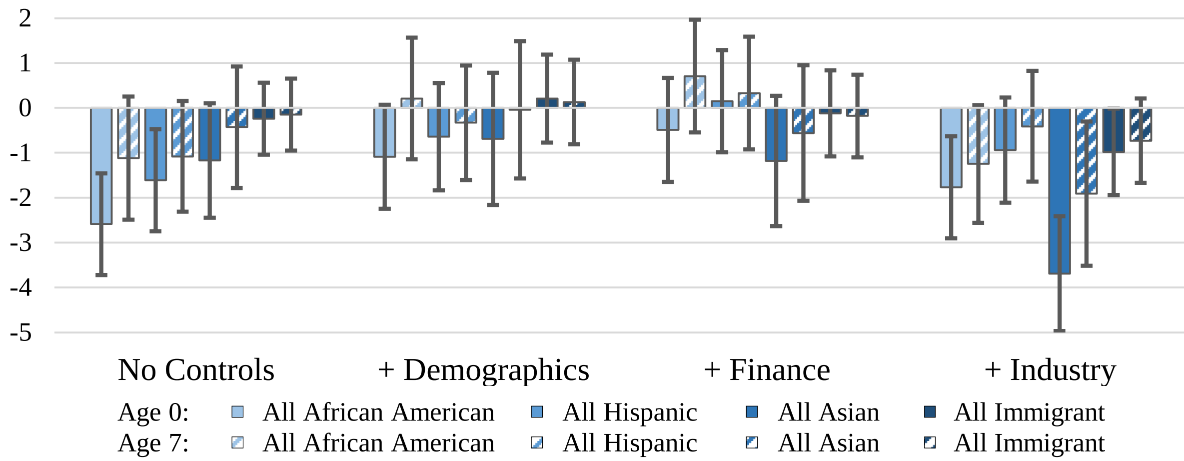
[22] Cabral and Mata (2003) also find firm size correlated with owner age at firm age 0 but not 7. They interpret this result as implying that younger entrepreneurs face higher financial constraints, which they reason should diminish over time, while skills should persist. Azoulay et al. (2018) report a positive association at firm age 5 of high growth and founder age. Neither of these studies is able to control for the amount of finance, however.

[23] Cabral and Mata suggest an alternative explanation for the diminished owner age effect over time, namely that firm-specific experience eventually overtakes previous owner labor market experience in importance. The attenuation of the age effect could also reflect the aging of the owners over seven years (many of them would be in the next higher age category if measured at firm age 7), but these factors alone would not account for the profile becoming negative.

growth entrepreneurships, so that traits such as flexibility and creativity may dominate labor market experience. The estimated owner age effects are not sensitive to industry selection.

Differences in high-growth performance by race, ethnicity, and citizenship could potentially result from discrimination in financial or product markets, as well as from correlated skills or preferences of individuals selecting into entrepreneurship. Previous research reports

Figure 4. High Growth Entrepreneurship by Race, Ethnicity, and Immigrant



Note: The omitted category for Race/Ethnicity is all non-Hispanic white, and additional variables not shown include other minorities and minority diversity (results in Figure 8 below). Full results are reported in Table B2.

significant differences in average business size and growth along these dimensions, but does not analyze the probability of high growth or large size.[24] Our results in Figure 4 (Appendix Table B3) imply lower rates of high-growth among Hispanic, African-American, and Asian-owned businesses, relative to all-white owned businesses, in most specifications. Most of the coefficients for these categories are statistically insignificant, however, with the exception of African Americans at firm age 0, who have a coefficient of -2.6 with no controls and -1.1 with demographic and human capital controls. The former is similar in magnitude to the gender gap at age 0, implying a 50 percent lower probability of operating a high-growth entrant (relative to a baseline of 5 percent), while the latter implies a 22 percent difference, and is only marginally significant. While the gender gap changed little from firm age 0 to 7, however, the African-American gap falls considerably and even becomes positive (but statistically insignificant) at age 7 in specifications with demographics and finance controls. Immigrant coefficients are usually small and statistically insignificant. The inclusion of start-up finance moves the Hispanic and African-American coefficients in a positive direction, consistent with these groups using less start-up capital. Including finance has the opposite effects on the Asian and immigrant coefficients, consistent with them using more start-up capital, but none are statistically significant. The coefficients for all these ethnicity and race categories move sharply in the negative direction with industry controls,

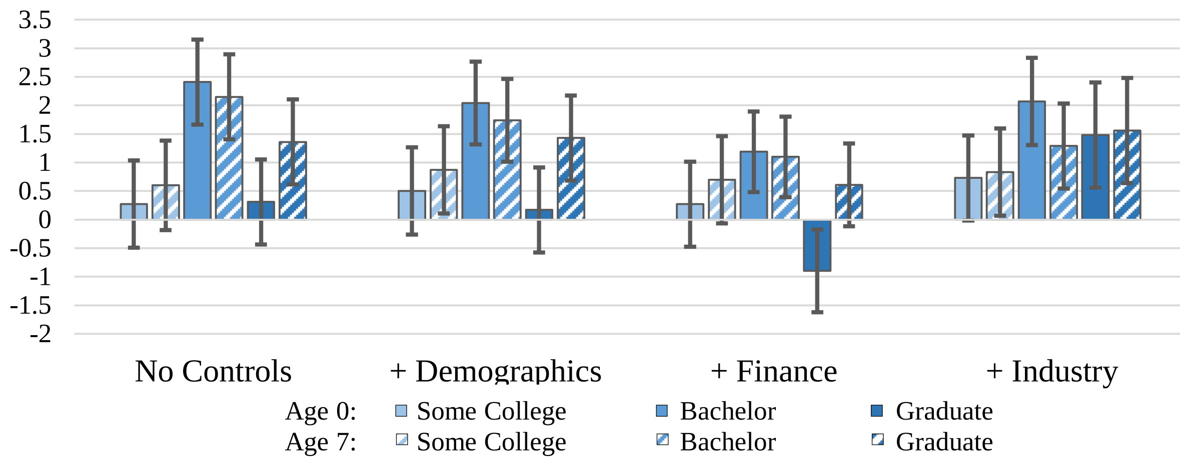
[24] Fairlie and Robb (2007, 2009) and Robb and Fairlie (2009) find that Native American-owned and Asian-owned businesses have higher average sales than White-owned businesses, while those of AfricanAmerican-owned businesses are lower. Kerr and Kerr (2017) find that immigrant-owned firms in the LBD start with lower average employment, not controlling for other owner or firm characteristics. Jarmin and Krizan (2010) find that Hispanic, African American, and other minorities (except Asian) have lower employment growth rates.

implying that minorities and immigrants are more prevalent in industries with higher shares of high-growth entrepreneurship.

Turning to measureable skills, formal education may increase an entrepreneur's ability to make decisions about business development. It may also be associated with better social networks and higher earnings prior to starting the business, increasing access to start-up capital.[25] Cabral and Mata (2003) report that owner education is positively associated with firm size at both age 0 and 7, but more strongly so at age 7. Our results in Figure 5 show coefficients for three types of education relative to high school graduate: some college, bachelor's degree, and graduate degree. Other educational categories (including less than high school, vocational and associate degrees) have slightly lower probabilities than high school graduates of being in the high-growth group, which is low for all of them (as shown in the full results in Table B4).

Entrepreneurs with bachelor's degrees have by far the highest probability of being in the top 5 percent, high-growth group. Their probability is nearly 50 percent of the baseline higher than

Figure 5. High Growth Entrepreneurship by Education



Note: The omitted category for founder education is high school diploma or GED, and additional variables not shown are less than high school, vocational, and associate degree. Full results are reported in Table B3.

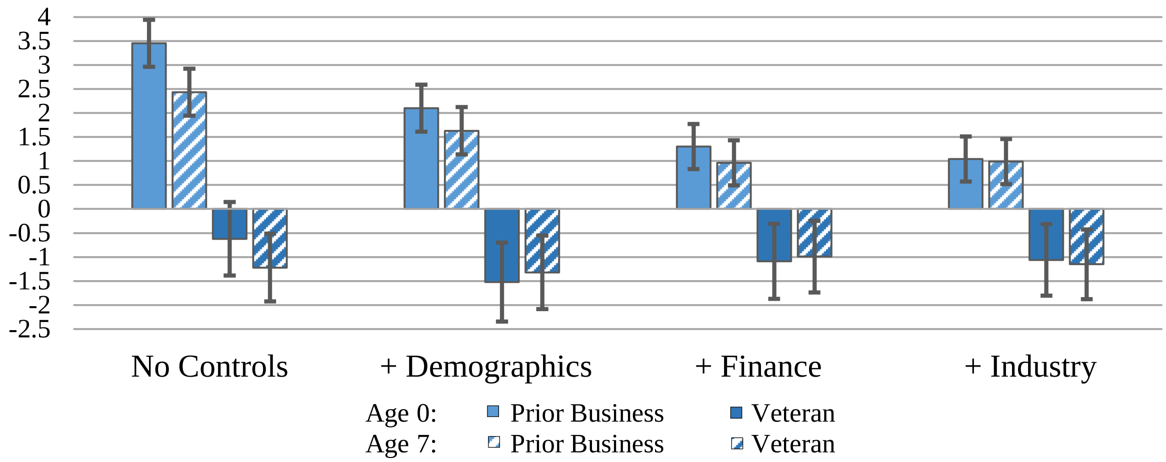
high school graduates in the specification with no controls. Those with graduate degrees, however, differ little from those with high school at age 0, implying a concave effect of education. The impact of graduate degree is much higher at age 7 than age 0, though, implying that a graduate degree is associated with high growth after start-up. These results hold up when controlling for other owner demographic characteristics, as also shown in Figure 5. Controlling for start-up finance lowers the bachelor and graduate degree coefficients, even leading to a negative coefficient for graduate degree recipients at the time of firm entry. Industry controls sharply raise the coefficients for graduate degree, and also for bachelor's degree at age 0. This means more highly-educated owners use more start-up capital and choose industries with a lower share of high-growth entrepreneurship. The latter result is consistent with Hurst and Pugsley's (2017) observation that skilled professions (e.g., dentists, doctors, lawyers, accountants, and insurance agents) are industries dominated by small businesses both when firms are young and old.

[25] See Baptista, Karaoz, and Mendonca (2014) for a discussion.

Two other types of human capital that might influence entrepreneurial performance are also measurable in our data: military service and previous entrepreneurial experience. There has been much policy interest in veterans, but there has been little study of their entrepreneurship. Using the 2002 SBO, Headd and Saade (2008) find that the size and industry distributions and of veteran-owned and non-veteran-owned firms are similar, without controlling for other factors, and they show that veterans and non-veterans have similar propensities to use different start-up financing sources. In the start-up cohort we analyze 9 percent of firms are owned by veterans, and as shown in Figure 6, they have a lower propensity to own firms in the high-growth entrepreneurship group at both age 0 and 7, with or without controls. The smaller magnitudes of coefficients with start-up finance controls are consistent with veterans using less start-up capital.

Previous experience in entrepreneurship may furnish owners with better managerial and

Figure 6. High Growth Entrepreneurship by Prior Business and Military Experience

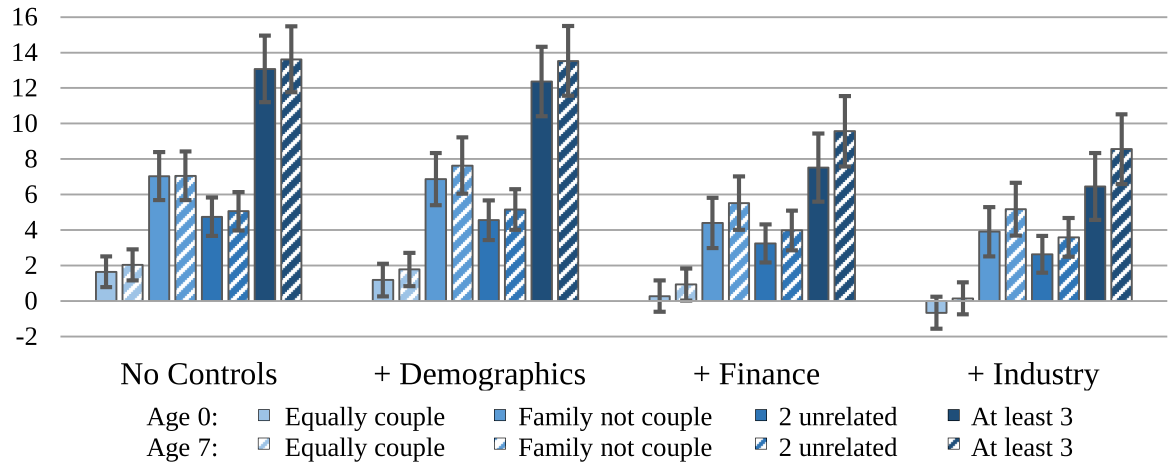


Note: Full results are reported in Table B3.

technical skills, a more developed business network, and greater knowledge about business opportunities (e.g., Baptista, Karaoz, and Mendonca 2014; Shaw and Sorenson 2017). It could also increase start-up capital via personal wealth accumulation, credit and entrepreneurial performance history, and an investor network. Shaw and Sorensen (2017) find that firms owned by serial entrepreneurs in Danish data have higher employment than those with novice entrepreneurs, but this result reverses once controlling for other owner and firm characteristics. We find that entrepreneurs with prior business ownership experience account for just over half of the founders in the sample, and as shown in Figure 6, they are twice as likely to be classified as high growth at age 0 and 63 percent more likely at age 7, compared to those with no prior ownership experience. With the baseline demographic, human capital, and founding team controls, these differentials decline but they are still substantial, as shown in Table 3.3: 40 percent at age 0 and 30 percent at age 7, both statistically significant. Controlling for the amount of start-up capital reduces both estimates to a little over 20 percent, consistent with serial entrepreneurs using more start-up capital, which Shaw and Sorensen (2017) also find. Adding detailed industry has little effect, implying that serial entrepreneurs are no more likely than novices to choose high-growth sectors.

The final set of issues concerns firms with multiple founders. About half the firms in our sample are founded by teams rather than single entrepreneurs, and the data permit us to investigate a number of interesting questions about the size and composition of the teams. A larger founding team can involve more diverse skill sets, providing a "jack of all trades" in a group that may be hard to find in an individual entrepreneur. More team members may also provide greater resources and networks for start-up capital.[26] The data in Table A1 show that nearly half the firms have multiple owners, with most of these (nearly 90 percent) being two-owner businesses. Start-ups are frequently family-owned, and the table implies that more than 70 percent of two-owner businesses are founded by related individuals, most of them married couples. Clearly, resources and interpersonal dynamics may differ in family and non-family teams.[27]

Figure 7. High Growth Entrepreneurship by Size and Family Composition of Founding Team



Note: The omitted owner type is one owner, and additional variables not shown represent firms owned primarily by husband and primarily by wife. Full results are reported in Table B4.

Figure 7 provides estimates of the high-growth probability for several types of teams compared to single owners: equally operated by married couple, non-couple family, two unrelated owners, and three or more unrelated owners. Table B5 shows full results for these categories, plus married couples where the husband is the predominant owners and where the wife is predominant. The probabilities for these later two types are insignificantly different from single owners. As the Figure makes clear, firms with three or more unrelated owners have by far the highest propensity to be in the group of high-growth entrepreneurship at both age 0 and 7. The raw difference and the coefficient controlling for demographics lie in the range 12-14, implying a 250 percent greater propensity compared with single owners. The next highest propensity is family businesses other than husband-wife, followed by two unrelated owners, husband-wife firms, firms primarily run by husband or wife, and single-founder firms. Estimates with controls for finance and industry preserves the ordering but reduces the differences across these groups. Larger founding teams evidently have a greater propensity to locate in industries with more high-growth entrepreneurship.

---

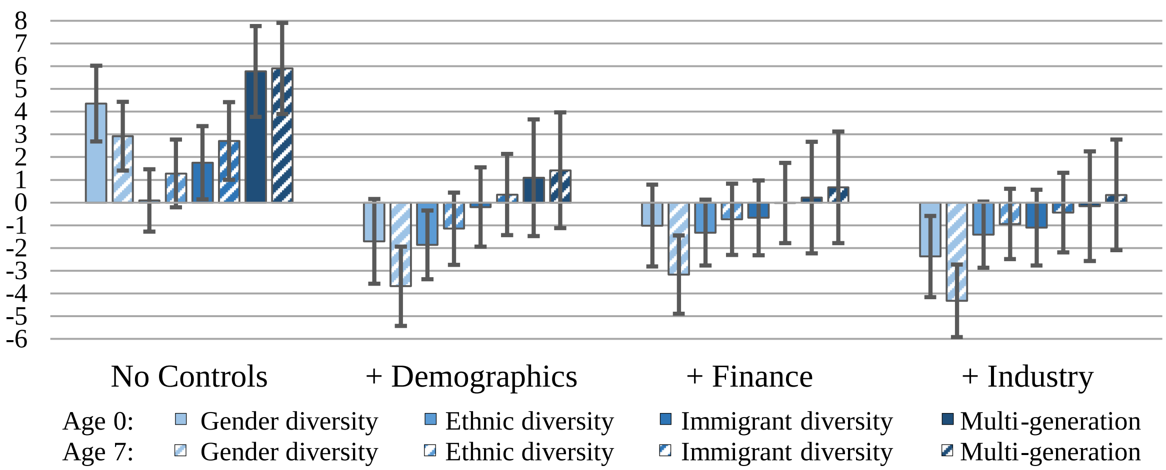
[26] See Ruef, Aldrich, and Carter (2003) for a discussion of the literature about founding teams. Baptista, Karaoz, and Mendonca (2014) find that firms with multiple owners have higher survival rates than single-owner firms.

[27] Brannon, Wiklund, and Haynie (2013) suggest that trust and familiarity are more important for a family business, while unrelated team members may be chosen based on skills and knowledge. They hypothesize that couples have worked out joint decision-making processes (e.g., about household finances), whereas non-couple family members are more likely to be in conflict with one another due to long-standing family roles, and their analysis of 295 teams from the PSED shows that couple-owned firms have a higher probability of ever having sales than other family firms.

Figure 8 contains estimates for variables representing four types of diversity within founding teams: by age, gender, race/ethnicity, and immigration status. A priori, it is unclear how diversity or similarity affect the high-growth probability. Similar founders may have easier communication, coordination, and trust-building. But diversity may imply varied skill sets and knowledge, leading to greater creativity and innovation, and may combine disparate traits in a team more easily than in single individuals, thus providing a team "jack of all trades" (Lazear 2004, 2005).[28]

The results for the specification with no controls show higher probabilities of high growth for all four types of diversity, relative to homogenous ownership, but once demographic controls are added the coefficients become either negative or statistically insignificant. The crucial difference across these specifications is that the demographic controls include team size. The estimated coefficients change only slightly when controls for start-up finance and industry are added. The clearest difference is that, relative to all male firms, those founded by gender-diverse teams are 60-80 percent less likely to become high growth.

Figure 8. High Growth Entrepreneurship by Diversity within Founding Team



Note: Full results are reported in Table B4.

4.3 Robustness

Research on the determinants of firm growth is marked by inconsistencies of results across studies. McKelvie and Wiklund (2010) point to several measurement issues potentially contributing to the inconsistencies, such as when in the lifecycle to begin tracking the firm, organic vs. acquisitive growth, and how the growth is measured. Growth factors could also vary with macroeconomic conditions, which may be particularly relevant for the sample in this study, given that the Great Recession began soon after the firms in our sample started up.

[28] Hoogendoorn and van Praag (2012) report that business performance decreases with increasing ethnic diversity below a certain share of minorities on the founding, team, but it becomes positive above a certain threshold. Hoogendoorn, Oosterbeek, and van Praag (2013) find that equally balanced male-female founding teams achieve higher profits than male-dominated teams.

To assess the importance of these and other concerns, we have conducted robustness exercises along several dimensions. As Shane (2008) notes, firms may need time to complete their initial hiring process, in which case employment in the first quarter of life may not be the right time to measure start-up size. When replacing age 0 employment with age 1 employment (four quarters after birth) in the employment transition matrix, we find very similar patterns, as shown in Appendix Tables C1-C2. Table C1 shows that firms starting large generally either stay large or exit, firms starting large have a much higher propensity to be large at age 7 than firms starting smaller, cohort employment is highly concentrated in large firms at birth, and the concentration is even higher at age 7. The employment transition matrix between age 0 and age 1 in Table C2 shows that firms that are large at age 0 make up the bulk of firms that are large at age 1 (60.5 percent), and the share of firms that are below 20 employees at age 0 that grow to 20 or more by age 1 is minor, for instance only 12 percent of firms with 10-19 employees at age 0 do so.

The patterns of association of founder characteristics with employment size are also very similar at age 0 and age 1. As shown in the regression results in Table C3, when replacing the top 5 percent of employment at age 0 dummy with that at age 1, differences in results are nearly all statistically insignificant. The only exceptions are coefficients that are still negative, but larger in magnitude, on all immigrant and on started the business more than two years before.

Another potential concern is the choice of start-up year, which in the case of 2007 immediately precedes the Great Recession. To investigate the extent to which the results are sensitive to this choice, we run regressions using the only other data containing all the same variables, from the 2012 SBO. Table C4 shows results for the four 2012 quarterly start-up cohorts, constructed according to the same procedures as described above for 2007. Most patterns are quite similar for the 2012 and 2007 cohorts at age 0 (of course we cannot yet examine the 2012 cohort at age 7). Comparing the 2012 relative to the 2007 cohort, there are a few statistically significant differences: the coefficients on all immigrant and primarily operated by the wife are negative and significant in 2012; and the coefficients on two unrelated owners, at least three unrelated owners, and start-up capital 1m and more are still positive but smaller in magnitude.

Related to the start-up year definition, we follow most previous research in identifying business entry as initial hiring (in the BR, the first quarter with paid employees), but we have also taken into account information on activities prior to hiring, in two different ways. First, the SBO asks each owner to report when the business was "established" (with no further definition). Second, we use information from the Census Bureau on businesses reporting revenue but no employees. From each of these, we construct variables capturing activity prior to hiring. Including these variables in the regressions has only negligible impact on the other results.

Multicollinearity is another issue for interpreting the results. The unconditional results with no controls have broadly similar patterns to those with demographic controls, but there are some differences as noted above. A more important issue is the possibility that some variables are jointly determined with firm size and growth, and for this reason we present the base specification that excludes financing and industry and show that indeed other variables have different estimated coefficients depending on whether financing and industry are included. Among the variables in the base specification, we identify founding team variables and time between initial start-up and hiring the first employee as those most likely to be problematic. We estimate regressions excluding these variables from the base specification and find qualitatively similar results.[29]

We also worried that results might be sensitive to the choice of the threshold for defining high growth. The estimates presented above use a top 5 percent threshold, but we have also run regressions using top 2 percent and top 10 percent thresholds, with results available upon request. The general patterns are very similar, but the magnitudes of several of the effects are monotonically increasing as the threshold rises from the top 10 to 5 to 2 percent, including the negative all female (at firm age 0), veteran, gender diversity, and ethnic diversity (at firm age 0) effects and the positive owner age (at firm age 0), bachelor's and graduate degree, prior business, two and three or more unrelated owners, and start-up finance effects.

A related issue is that our analysis includes exiting firms in the calculation of the top 5 percent at age 7, but because of the high exit rate, the firms in the top 2 percent of this distribution at age 7 are roughly the same as the top 5 percent of survivors.

Other robustness checks include the following: We have estimated logistic regressions in place of linear probability models. Since there is some question whether firms with very high employment when they first appear are really start-ups, we have run regressions dropping all firms with 100 or more employees in their first quarter.[30] We have estimated regressions without adjusting for boundary changes. We have also estimated regressions with owner share weights, but not LBD weights, to examine the sensitivity of the results to LBD weighting, and we have added franchises to the sample. The results from all of these variants are qualitatively similar to those reported above, and they are available upon request.

5. Conclusion

Building on previous research reporting the predominance of job creation among entrants and young firms, this paper has analyzed high-growth entrepreneurship using a unique data set. The data are based on a large, random, and nationally representative sample of firms that permits us to follow an entry cohort, dealing with the challenges of defining when firms start up and tracking them over time. The data we have constructed are also unusual in containing information immediately after the firm starts up on a rich set of founder characteristics.

Our empirical results confirm the finding in previous research of large skewness in the employment size distribution, whereby a small fraction of firms account for most employment in the U.S. economy. We add to this result by documenting the skewness on entry and 7 years later, showing the importance of high-growth entrepreneurship at both ages and the high persistence of size over the first 7 years of a firm's life. We show that large entrants are not only much more likely to be large firms at age 7, but they are also more dynamic than their smaller counterparts: among growing firms, the average job creation is largest among large entrants, while among contracting and exiting firms they destroy the most jobs on average. Thus, the "up or out dynamic" of Decker et al. (2014) is strongest for this group. By contrast, small entrants are most likely to remain small and not grow, consistent with Hurst and Pugsley (2011). Exit rates are high across

[29] Since we omit founding team variables like couples and different types of diversity in these specifications, we control for owner-level gender, ethnicity, race, and immigrant variables rather than firm-level variables indicating whether all of the owners are in the particular category.

[30] Firms with 100 or more employees are 0.5 percent of the firms in the sample in Tables 2 and 3.

the board, only declining moderately with start-up size. We also find that most of the size variation is within rather than between narrow (6-digit) industries and little is explained by differences in the amount of start-up capital.

These findings suggest that the upper tail of the entrant size distribution holds particular interest and motivate our detailed examination of founder characteristics that may predict high growth, here defined as the top ventile of the size distribution at age 0 and age 7. Controlling in a base specification for other aspects of demographic, human capital, and founding team characteristics, but not for start-up finance and industry, we find a much lower probability (34 percent) for women to operate a high-growth business. This large gender gap falls to about 20 percent when we control for the amount of start-up finance, which could result from discrimination in financial access or from non-pecuniary motives for founding the business, neither of which we can observe directly. But the gap rises to about 40 percent when we add detailed industry effects, which suggests that women-owned businesses tend to be disproportionately in high-growth sectors. The estimated effects are similar at age 0 and age 7.

By contrast with the large and statistically significant gender gap, we generally find only modest differences with respect to race, ethnicity, and nationality. Unconditional estimates imply a lower probability of operating a high-growth firm for Hispanic, African-American, and Asian owners, but coefficients for these categories are generally insignificant when we control for other variables in the base specification. The major exception is African-American owners at firm age 0, who have a 28 percent lower probability of entering in the largest 5 percent of businesses, conditional on other demographic and skill characteristics, but this difference disappears by firm age 7. When the amount of start-up capital is controlled the racial gap also falls and becomes statistically insignificant at age 0, and it moves in a positive direction at age 7, a result which is consistent with financial discrimination at start-up. But it rises in magnitude (and remains negative) when detailed industry controls are added, implying that African Americans select into industries with higher shares of large entrants. We find no significant differences for immigrants in any specification, a finding at odds with some popular beliefs and prior research on immigrant entrepreneurship.

Concerning the age of the founder(s), the high-growth probability is clearly positively sloped at firm age 0, but then flattens by the time firms reach age 7. The positive profile at age 0 might be explained by lower skills or greater financial constraints faced by younger entrepreneurs, and consistent with the latter we find that controlling for start-up finance yields a flatter founder age profile at that firm age. The flat profile in the base specification at age 7 might be explained by leveling out of the financial constraints as the firm ages, but when we control for start-up finance, the slope of the founder age profile becomes negative, implying that the effect of any tougher financial constraints for younger entrepreneurs tends to persist. A negative slope is inconsistent with general labor market experience playing an important role in entrepreneurial human capital.

With respect to formal education, a striking finding at firm age 0 is that bachelor’s degrees are associated with a much higher (31 percent) probability of large size than either high school or graduate degrees. The difference vis-à-vis high school largely persists to age 7, but it becomes negligible with respect to graduate education. Controlling for start-up finance actually yields a significantly lower probability of high growth for those with graduate versus only high-school education, although again this disappears 7 years later. Controlling for detailed industry raises the graduate education coefficient, implying that this group tends to choose sectors with relatively small firms, perhaps because many of them work in professions such as law, medicine, or accounting. In any case, our results do not support an important role for graduate education in producing high-growth entrepreneurs. Perhaps less surprising, we find that military experience is negatively associated and prior business ownership is positively associated with high growth, in both cases strong results that are robust over time and across specifications.

Finally, concerning founding teams we find that businesses with more founders are more likely to be large at start-up and subsequently. The differences are larger than any others in the data: in the baseline specification, firms with at least three unrelated founders are 230 percent more likely and those with two founders other than a couple are about 100 percent more likely to be in the top ventile than single owners. Diversity in age, race/ethnicity, and immigration status have little association with the high growth probability, while gender diversity is negatively associated.

Some of these results differ from patterns previously reported by other researchers in related research. The differences may be explained by sample sizes and representativeness, focus on averages versus high growth, definitions of high growth, or analyzing a cross-section versus following an entry cohort. In other cases, the results in this paper confirm previous research, putting them on a more secure footing. A series of robustness checks confirm that our findings are not sensitive to small changes in specification or in the definition of high growth.

We find not only that firm size is highly persistent from entry to age 7, but also so are the relationships of high growth with characteristics. There is a general tendency for an attenuation of coefficients from age 0 to age 7, suggesting increasing difficulty in accounting for growth heterogeneity. But except in a few cases we have noted, the qualitative patterns are similar, and most of the differences are statistically insignificantly different from zero at conventional levels. Thus, the patterns observed at age 0 already embody most of what one can learn 7 years later.

On the other hand, estimates frequently vary with controls for the amount of start-up finance and for the industry in which the firm operates. We have argued that these variables are particularly suspect for the possibility of correlation with important unobservables such as motivations and skills of the entrepreneur and the quality of the business idea, so we have excluded them from our base specifications, but in several cases adding them to a richer specification is helpful in illuminating the patterns with respect to other variables. Future research could fruitfully focus on these relationships.

References

Acs, Zoltan J., William Parsons and Spencer Tracy, 2008, "High-Impact Firms: Gazelles Revisited," Washington D.C.: Small Business Administration.

Angelini, Paolo, and Andrea Generale, 2008, "On the Evolution of Firm Size Distributions," *American Economic Review*, 98(1), 426-38.

Azoulay, Pierre, Benjamin F. Jones, J. Daniel Kim, and Javier Miranda, 2018, "Age and High-Growth Entrepreneurship," NBER Working Paper No. 24489.

Baptista, Rui, Murat Karaoz, and Mendonca, 2014, "The Impact of Human Capital on Early Success of Necessity vs. Opportunity-Based Entrepreneurs," *Small Business Economics*, 42(4), 831-847.

Bates, Timothy, 1990, "Entrepreneur Human Capital Inputs and Small Business Longevity," *Review of Economics and Statistics*, 72(4), 551-559.

Birch, David L., 1979, *The Job Generation Process*, unpublished report prepared by the MIT Program on Neighborhood and Regional Change for the Economic Development Administration, U.S. Department of Commerce, Washington, DC.

Birch, David L., 1981, "Who Creates Jobs?" *The Public Interest*, 65, 3-14.

Birch, David L., 1987, *Job Creation in America: How Our Smallest Companies Put the Most People to Work*, Free Press, New York.

Blanchard, Lloyd, Bo Zhao, and John Yinger, 2008, "Do Lenders Discriminate against Minority and Woman Entrepreneurs?" *Journal of Urban Economics*, 63(2), 467–97.

Blanchflower, David G., Phillip B. Levine, and David J. Zimmerman, 2003, "Discrimination in the Small-Business Credit Market," *The Review of Economics and Statistics* 85(4), 930–43.

Bloom, Nick, and John Van Reenen, 2007, "Measuring and Explaining Management Practices Across Firms and Countries," *Quarterly Journal of Economics*, 122(4), 1351-1408.

Brannon, David L., Johan Wiklund, and J. Michael Haynie, 2013, "The Varying Effects of Family Relationships in Entrepreneurial Teams," *Entrepreneurship Theory and Practice*, 37(1), 107-132.

Brown, J. David, and John S. Earle, 2017, "Finance and Growth at the Firm-Level: Evidence from SBA Loans," *Journal of Finance,* 72, 1039–1080.

Brown, J. David, John S. Earle, and Dana Lup, 2005, "What Makes Small Firms Grow? Finance, Human Capital, Technical Assistance, and the Business Environment in Romania," *Economic Development and Cultural Change*, 54(1), 33-70.

Cabral, Luís M. B., and José Mata, 2003, "On the Evolution of the Firm Size Distribution: Facts and Theory," *American Economic Review*, 93(4), 1075–90.

Coad, Alex, Julian S. Frankish, Paul Nightingale, and Richard G. Roberts, 2014, "Business Experience and Start-up Size: Buying More Lottery Tickets Next Time Around?" *Small Business Economics*, 43, 529-547.

Coleman, Susan, and Alicia Robb, 2014, "Access to Capital by High-Growth Women-Owned Businesses," report SBAHQ-13-Q-0A63 for the National Women's Business Council.

Daunfeldt, Sven-Olov, and Daniel Halvarsson, 2015, “Are High-Growth Firms One-Hit Wonders? Evidence from Sweden,” *Small Business Economics,* 44, 361-383.

Daunfeldt, Sven-Olov, Dan Johansson, and Daniel Halvarsson, 2015, “Using the Eurostat-OECD Definition of High-Growth Firms: A Cautionary Note,” *Journal of Entrepreneurship and Public Policy*, 4(1), 50-56.

Decker, Ryan, John Haltiwanger, Ron Jarmin, and Javier Miranda, 2014, “The Role of Entrepreneurship in U.S. Job Creation and Economic Dynamism,” *Journal of Economic Perspectives*, 28(3), 3-24.

Eurostat-OECD, 2008, *Eurostat-OECD Manual on Business Demography Statistics*, OECD Publishing: Paris, France.

Fairlie, Robert W., and Javier Miranda, 2017, “Taking the Leap: The Determinants of Entrepreneurs Hiring Their First Employee,” *Journal of Economics and Management Strategy*, 26(1), 3-34.

Fairlie, Robert W., and Alicia M. Robb, 2007, “Why Are Black-Owned Businesses Less Successful than White-Owned Businesses? The Role of Families, Inheritances, and Business Human Capital,” *Journal of Labor Economics*, 25(2), 289–323.

Fairlie, Robert W., and Alicia M. Robb, 2009, “Gender Differences in Business Performance: Evidence from the Characteristics of Business Owners Survey,” *Small Business Economics*, 33(4), 375-395.

Fairlie, Robert, Alicia Robb, and David T. Robinson, 2016, *Black and White: Access to Capital among Minority-Owned Startups*, SIEPR Working Paper No. 17-03.

Foster, Lucia, Cheryl Grim, and John Haltiwanger, 2016, “Reallocation in the Great Recession: Cleansing or Not?” *Journal of Labor Economics*, 34(1), S293-S331.

Frank, Murray Z. 1988. “An Intertemporal Model of Industrial Exit,” *The Quarterly Journal of Economics*, 103(2), 333–344.

Garnsey, Elizabeth, Erik Stam, and Paul Heffernan, 2006, “New Firm Growth: Exploring Processes and Paths,” *Industry and Innovation*, 13(1), 1-20.

Haltiwanger, John, Ron S. Jarmin, and Javier Miranda, 2013, “Who Creates Jobs? Small versus Large versus Young,” *Review of Economics and Statistics*, 95(2), 347–361.

Headd, Brian, and Radwan Saade, 2008, “Do Business Definition Decisions Distort Small Business Research Results?” SBA Office of Advocacy Working Paper.

Holzl, Werner, 2014, “Persistence, Survival and Growth: A Closer Look at 20 Years of Fast-Growing Firms in Austria,” *Industrial and Corporate Change*, 12(1), 199-231.

Hoogendoorn Sander, and Mirjam van Praag, 2012, “Ethnic Diversity and Team Performance: A Field Experiment,” Tinbergen Institute Discussion Paper TI 2012-068/3.

Hoogendoorn, Sander, Hessel Oosterbeek, and Mirjam van Praag, 2013, “The Impact of Gender Diversity on the Performance of Business Teams: Evidence from a Field Experiment,” *Management Science*, 59(7).

Hopenhayn, Hugo A., 1992, “Entry, Exit, and Firm Dynamics in Long Run Equilibrium,” *Econometrica*, 60 (5): 1127–1150.

Hurst, Erik, and Benjamin Wild Pugsley, 2011, "What Do Small Businesses Do?" *Brookings Papers on Economic Activity*, Fall 2011, 73-118.

Hurst, Erik G., and Benjamin W. Pugsley, 2017, "Wealth, Tastes, and Entrepreneurial Choice," in *Measuring Entrepreneurial Businesses: Current Knowledge and Challenges*, John Haltiwanger, Erik Hurst, Javier Miranda, and Antoinette Schoar, eds.

Jarmin, Ron, and C.J. Krizan, 2010, "Past Experience and Future Success: New Evidence on Owner Characteristics and Firm Performance," CES Discussion Paper 10-24.

Jarmin, Ron, C.J. Krizan, and Adela Luque, 2014, "Owner Characteristics and Firm Performance During the Great Recession," CES Discussion Paper 14-36.

Jovanovic, Boyan, 1982, "Selection and the Evolution of Industry," *Econometrica*, 50(3), 649–670.

Kalleberg, Arne L., and Kevin T. Leicht, 1991, "Gender and Organizational Performance: Determinants of Small Business Survival and Success," *The Academy of Management Journal*, 34(1), 136–161.

Kerr, Sari Pekkala, and William R. Kerr, 2017, "Immigrant Entrepreneurship," in *Measuring Entrepreneurial Businesses: Current Knowledge and Challenges*, John Haltiwanger, Erik Hurst, Javier Miranda, and Antoinette Schoar, eds., University of Chicago Press.

Lazear, Edward P., 2004, "Balanced Skills and Entrepreneurship," *American Economic Review*, 94, 208-211.

Lazear Edward P., 2005, "Entrepreneurship," *Journal of Labor Economics*, 23, 649-680.

Lotti, Francesca, and Enrico Santarelli, 2004, "Industry Dynamics and the Distribution of Firm Sizes: A Nonparametric Approach." *Southern Economic Journal*, 70(3), 443–66.

McKelvie, Alexander, and Johan Wiklund, 2010, "Advancing Firm Growth Research: A Focus on Growth Mode Instead of Growth Rate," *Entrepreneurship Theory and Practice*, 34(2), 261–288.

Melitz, Marc J., 2003, "The Impact of Trade on Intra-Industry Reallocations and Aggregate Industry Productivity," *Econometrica*, 71(6), 1695–1725.

Neumark, David, Brandon Wall, and Junfu Zhang, 2011, "Do Small Businesses Create More Jobs? New Evidence for the United States from the National Establishment Time Series," *Review of Economics and Statistics*, 93(1), 16-29.

Nightingale, Paul, and Alex Coad, 2014, "Muppets and Gazelles: Political and Methodological Biases in Entrepreneurship Research," *Industrial and Corporate Change*, 23(1), 113-143.

OECD, 2010, *High-Growth Enterprises: What Governments Can Do to Make a Difference*, OECD Studies on SMEs and Entrepreneurship, OECD Publishing.

Robb, Alicia M., and Robert W. Fairlie, 2009, "Determinants of Business Success: An Examination of Asian-Owned Businesses in the US," *Journal of Population Economics*, 22, 827-858.

Ruef, Marin, Howard E. Aldrich, and Nancy M. Carter, 2003, "The Structure of Founding Teams: Homophily, Strong Ties, and Isolation Among U.S. Entrepreneurs," *American Sociological Review*, 68(2), 195-222.

Shane, Scott, 2008, *The Illusions of Entrepreneurship: The Costly Myths That Entrepreneurs, Investors, and Policy Makers Live By*, New Haven: Yale University Press.

Shaw, Kathryn L., and Anders Sorensen, 2017, "The Productivity Advantage of Serial Entrepreneurs," NBER Working Paper No. 23320.

Stangler, Dane, 2010, "High-Growth Firms and the Future of the American Economy," Kauffman Foundation Research Series: Firm Formation and Economic Growth.

Storey, D. J., 1994, "New Firm Growth and Bank Financing," *Small Business Economics*, 6(2), 139-150.

Syverson, Chad, 2011, "What Determines Productivity?" *Journal of Economic Literature*, 49(2), 326-365.

U.S. Census Bureau, 2016, Business Dynamics Statistics: Firm Characteristics Data Tables – Firm Age, https://www.census.gov/ces/dataproducts/bds/data_firm.html, accessed on September 6, 2017.

Appendix A: Data

In order to measure the precise level of employment at start-up, track firm employment over time, and incorporate characteristics of business founders, we link together several data sources. We start with the U.S. Census Bureau's Business Register (BR), which includes all nonfarm businesses filing Internal Revenue Service tax forms as individual proprietorship, partnership, or any type of corporation, and with receipts of $1,000 or more. The BR is available quarterly, and employment is the number of employees in the payroll period including March 12 for quarter 1, June 12 (quarter 2), September 12 (quarter 3), and December 12 (quarter 4), as reported to the Internal Revenue Service at the Employer Identification Number (EIN) level. Different units within a firm may file under separate EINs each quarter, and we aggregate such cases to obtain firm-level employment.

We define entry as hiring a first employee, so entrants in a particular quarter are firms with positive employment for the first time. We also combine information from the Longitudinal Business Database (LBD), an annual database containing all non-farm businesses with positive payroll in the year, and we require that an entrant first appear in the LBD the same year and that none of an entrant's establishments that year had positive employment beforehand in the LBD. We restrict the sample to firms found in the LBD because we use the LBD's longitudinal establishment links across annual BR files to track firms and their reorganizations over time. We take employment from the BR rather than the LBD, because the LBD only contains employment in the pay period including March 12. Age 0 is the firm's first quarter with positive employment.[31]

For detailed information on characteristics of firms and owners, we focus on the four quarterly entry cohorts in 2007 in order to link these data to the Census Bureau's 2007 Survey of Business Owners (SBO).[32] The SBO uses the BR as the sampling frame, stratified by state, industry, owner demographic group, and whether the firm has employees or not.[33] The largest companies in each stratum are selected with certainty, and the remainder of the sample is randomly selected. The SBO has been carried out every five years, and we use the 2007 SBO rather than the 2002 or 2012 data because the 2002 SBO lacks information on several of the factors we wish to study, and the 2012 SBO permits observation only on a short time span after start-up.[34] Motivated by concerns about whether results differ for firms founded in 2007, just before the Great Recession, compared to firms founded in other years, we have also estimated all the Age 0 relationships with

[31] As with any longitudinal data, there is a possibility of broken links, and the data do contain a small number of implausibly large entrants. But while such outliers may have large leverage on estimated effects in standard employment regressions, our approach of estimating the probability of being in a high-growth group gives no extra leverage to these firms and is therefore more robust to such measurement errors. We also find similar results when we exclude all observations over 100 employees, as discussed in the robustness subsection below.

[32] Choosing the most recent start-up cohort reduces but does not completely eliminate survival bias, as the survey was conducted in 2008 and 2009 for the 2007 reference year. This problem is larger for studies of the 1992 CBO, which was carried out in 1996 for reference year 1992.

[33] The SBO does not collect ownership information on firms without an individual owner of at least 10 percent or that are majority owned by another company or organization, Employee Stock Ownership Plan, members in a cooperative or club, an estate or trust, an Alaska Native Regional or Village Corporation, or an American Indian tribal entity.

[34] The 2002 SBO does not contain information on whether the owner was born in the U.S. or not, husband-wife ownership, prior business ownership, and amount of start-up capital.

the 2012 data. The results from this analysis, discussed in the robustness sub-section and provided in the Appendix, are very similar to those from 2007.

Tracking a firm longitudinally involves measurement problems that are much more prevalent among older and larger firms.[35] In particular, the firm identifier can change due to a re-registration, change in legal form, or switch from single- to multi-establishment. The LBD is designed to track such changes, but we implement additional procedures to restore broken firm links from identifier switches.[36] Approximately 6.3 percent of firms in the top 5 percent of employment at age 7 in the Table 2 and 3 sample below undergo a firm identifier switch between age 0 and age 7, vs. 2.9 percent of firms in the bottom 95 percent. This is consistent with growing firms being more likely to change legal form and/or become a multi-establishment firm.

Our focus is on the firm's organic growth, rather than growth through acquisition, so we adjust age 7 employment to remove the effects of establishment acquisitions and divestments. By definition of entry in our analysis, age 0 employment excludes acquisition of existing establishments. We use the LBD to track subsequent boundary changes with procedures similar to Haltiwanger, Jarmin, and Miranda (2013) and Brown and Earle (2017). If the firm acquires a pre-existing establishment, the establishment's employment in the year prior to acquisition is subtracted from the firm's age 7 employment.[37] If an establishment is sold or spun off and continues to operate in subsequent years, the establishment's employment in the year prior to divestment is added to the firm's age 7 employment. The reasoning behind these boundary change adjustment procedures is that the firm is responsible only for those establishment employment changes that occur while the establishment is under its control. In practice for the sample analyzed below these adjustments are inconsequential, because while a higher share of firms in the top 5 percent of employment have an employment adjustment due to boundary changes relative to the bottom 95 percent, the share is well below 1 percent in both groups. If a firm disappears from the LBD prior to 2014, and none of its establishments continue to operate subsequently, we treat it as an exit, imputing zero for age 7 employment.[38] In contrast, if at least one of its establishments in the firm's last year in the LBD continues to operate in subsequent years, we impute the firm's employment (or boundary-adjusted employment if it had boundary changes) in its last year in the LBD as age 7 employment.[39]

---

[35] McKelvie and Wiklund (2010) discuss challenges of tracking firms across time when measuring growth.

[36] We link firm identifiers A and B if identifier A last appears in year t and identifier B first appears in t+1, at least one establishment is in A in t and B in t+1, the establishments in A in t and B in t+1 (denoted A-B) have more total employment in t than any other establishment groups in A in t switching to some other firm identifier in t+1 (A-C, A-D, etc.), and A-B has more employment in t than any other establishment groups switching to B in t+1 from another firm identifier (E-B, F-B, etc.).

[37] The BR may sometimes misclassify new establishment openings by pre-existing firms as new firms. The Census Bureau learns about such establishment-firm linkages during the quinquennial economic census, but it does not know when the firm first owned the establishment. In such cases we misclassify new establishment openings as acquisitions and undercount the firm's organic growth. It is unlikely that any of the firms in our SBO sample are new establishments owned by pre-existing firms, because businesses majority owned by other businesses are not asked for owner information.

[38] Just as firms first appearing with pre-existing establishments are not classified as entrants, firms that disappear without closing down all their establishments are not classified as exits.

[39] In the sample for Tables 2 and 3, 7.3 percent of firms from the age 7 top 5 percent exit before age 7 with subsequently continuing establishments, compared to 0.2 percent of firms in the bottom 95 percent.

The sample used for the analysis of characteristics consists of all owner observations for firms in the four quarterly 2007 LBD start-up cohorts in the 2007 SBO that have non-missing values for all the characteristics. The sample size is 55,800 owners of 37,100 firms, which is about 7.0 percent of all firm start-ups that year (U.S. Census Bureau, 2016). To make the analysis firm-level, we weight each owner by the ownership equity shares (so they sum to one for each firm). To reflect the industry-size composition of the LBD, we also weight by the inverse of the sample-population ratio (the share of firms in the two-digit NAICS industry–employment category in the 2007 quarterly start-up cohorts in the LBD divided by the sample's share of firms in the two-digit NAICS industry–employment category). The size categories for these weights are 1, 2-4, 5-9, 10-19, and 20 and more employees. We use LBD-based weights for the 2007 start-up cohorts rather than SBO survey weights, because the SBO survey weights do not take nonresponse or firm age (a crucial variable for our analysis) into account.[40]

All the independent variables are measured for the year 2007, the start-up year for the firms in the main sample and the reference year in the SBO. Founder characteristics from the SBO include basic demographics (age, gender, race, ethnicity, and immigrant/native), human capital (type of education, veteran, and prior entrepreneurial experience), and the size of the founding team and relationships among multiple founders (family/unrelated and diversity by demographics and education). We use the firm's 6-digit industry from the BR, and categories of the amount of start-up finance from the SBO (0-5, 5-10, 10-25, 25-50, 50-100, 100-250, 250-1mln, >1mln, in 1000s of $; a small number respondents answered "unknown" or "no capital needed," which we control for but do not report).

Means of the finance variables and founder characteristics are provided in Appendix Table A1. Details of the construction of characteristics measures from the raw data are as follows. Among races, we distinguish whites, African Americans, and Asians, and we group native Hawaiians, Guamanian or Chamorro, Samoan, and other Pacific Islanders and some other race as "other minorities." Immigrant indicates the owner was not born in the United States. For firms with multiple owners, the gender, race, ethnicity, and immigrant variables are defined to indicate whether all the firm's owners are in that category or not, and thus include a label "all," in order to permit us to measure the impact of diversity, as discussed below. In the case of gender diversity, for instance, we define "all female" and "all male" variables to indicate firm with owners from only one gender or the other (including single owners).

Among human capital variables, the educational categories are self-explanatory. Veteran indicates whether the owner is a veteran of any branch of the U.S. military service, including the Coast Guard. Prior business indicates that the owner previously owned a different business prior to owning the current business.

The data permit us to construct detailed measures of the size and composition, including family relationships, of founding teams. We define diversity variables as follows: Gender diversity indicates that the business is jointly owned by at least one owner of each gender (except when husband and wife, for which we provide a separate category), ownership ethnic diversity indicates that the business is jointly owned by at least two individuals with different race or ethnicity from one another, ownership immigrant diversity is a dummy equal to 1 when the business is jointly

[40] Foster, Grim, and Haltiwanger (2016) apply similar LBD weights when using the Annual Survey of Manufacturers.

owned by individuals who are immigrants and U.S.-born, and multi-generation indicates that at least one owner is 20 or more years older than another. By controlling for all-female, the variables for different types of husband-wife ownership and gender diversity for non-couples measure whether gender ownership effects vary depending on who else co-owns the firm. Similarly, including variables for all of one ethnicity or race, all immigrant, and ethnic/racial and immigrant diversity in the regression allow us to examine whether race/ethnicity and immigrant effects differ with homophily or diversity among founder teams along those dimensions.

Table A1. Descriptive Statistics for Independent Variables

	% of Sample	% of High Growth			% of Sample	% of High Growth	
Variables	**Age 0**	**Age 0**	**Age 7**	**Variables**	**Age 0**	**Age 0**	**Age 7**
Start-up Finance				***Human Capital***			
No capital needed	6.9	4.9	4.1	Less than high school	3.5	2.8	2.0
Capital under 5k	21.1	7.0	6.7	High school	17.9	15.1	14.4
Capital 5k to 10k	9.7	3.0	3.0	Vocational school	5.7	4.1	4.0
Capital 10k to 25k	12.8	4.6	6.4	Some college	16.9	15.6	15.9
Capital 25k to 50k	9.8	4.6	7.2	Associate degree	6.1	4.3	5.0
Capital 50k to 100k	11.0	8.4	10.1	Bachelor's degree	29.2	38.8	36.2
Capital 100k to 250k	11.3	15.0	15.5	Graduate degree	19.9	18.1	21.4
Capital 250k to 1m	7.8	25.1	22.9	Veteran	9.0	8.1	7.1
Capital 1m and more	2.5	15.3	12.4	Non-veteran	91.0	91.9	92.9
Don’t know amount	7.0	12.1	11.6	Prior business	50.8	67.6	62.8
				No prior business	49.2	32.4	37.2
Demographics				***Founding Team***			
All female	16.0	8.2	8.5	One owner	50.6	29.3	28.1
All male	51.4	60.2	60.6	Equally by couple	12.1	11.0	11.6
Age <35	19.6	11.9	19.6	Primarily husband	10.8	7.4	7.4
Age 35-44	31.3	29.2	31.8	Primarily wife	4.3	3.1	3.2
Age 45-54	29.3	33.5	29.0	Family non-couple	8.3	16.4	16.2
Age 55-64	15.4	19.0	15.0	Two unrelated	10.5	15.9	16.3
Age 65 or over	4.0	5.9	4.0	At least 3 unrelated	5.9	18.5	19.1
All Hispanic	5.0	3.6	4.0	Multi-generation	2.9	6.1	6.2
All non-Hispanic	93.0	94.1	93.2	No multi-generation	97.1	93.9	93.8
All White	86.4	87.8	86.9	Gender diversity	5.4	10.1	8.7
All African American	2.8	1.5	2.2	No gender diversity	94.6	89.9	91.3
All Asian	4.8	3.7	4.3	Ethnic diversity	4.4	4.5	5.5
All other minority	3.3	4.4	3.6	No ethnic diversity	95.6	95.5	94.5
All immigrant	15.0	14.2	14.3	Immigrant diversity	4.3	5.7	6.5
All U.S.-Born	80.7	80.1	79.2	No immigrant diversity	95.7	94.3	93.5

Note: Source: 2007 SBO. “High growth” refers to the top ventile (5%) of the employment size distribution at age 0 and age 7, respectively.

Appendix B. Linear Probability Model (LPM) Results

Table B1. High Growth Entrepreneurs by Start-up Finance

	No Controls		Plus Demographics		Plus Industry	
Characteristic at Age 0	Top 5% Age 0	Top 5% Age 7	Top 5% Age 0	Top 5% Age 7	Top 5% Age 0	Top 5% Age 7
Capital 5k to 10k	-0.16	-0.08+	-0.21^	-0.23+^	-0.41	-0.34
	(0.28)	(0.27)	(0.28)	(0.27)	(0.29)	(0.28)
Capital 10k to 25k	0.06	0.85+	-0.25^	0.47+^	-0.69^	0.36+
	(0.25)	(0.29)	(0.25)	(0.29)	(0.28)	(0.30)
Capital 25k to 50k	0.68	2.09+	0.32^	1.55+^	-0.67^	1.12+^
	(0.36)	(0.41)	(0.35)	(0.41)	(0.37)	(0.41)
Capital 50k to 100k	2.15	2.97+	1.38^	2.15+^	-0.33^	1.22+^
	(0.40)	(0.44)	(0.40)	(0.44)	(0.42)	(0.46)
Capital 100k to 250k	5.07	5.29+	3.82^	3.92^	1.72^	2.81^
	(0.54)	(0.52)	(0.54)	(0.53)	(0.52)	(0.53)
Capital 250k to 1m	14.65	13.17	12.73^	11.09^	9.53^	9.28^
	(0.92)	(0.88)	(0.92)	(0.87)	(0.86)	(0.86)
Capital 1m and more	29.76	23.66+	26.22^	20.11+	22.57^	17.72+^
	(1.87)	(1.73)	(1.87)	(1.74)	(1.79)	(1.73)
Don't know amount	7.02	6.71	5.99^	5.80^	4.36^	4.96^
	(0.63)	(0.67)	(0.62)	(0.66)	(0.61)	(0.63)
No capital needed	1.89	1.35+	1.71^	1.34^	1.33^	0.98^
	(0.40)	(0.43)	(0.40)	(0.42)	(0.39)	(0.41)

Note: Results from LPM estimation of Equation (1). Coefficients and standard errors are multiplied by 100 for ease of reading. Each regression also includes the variables in Tables A4.1, A4.2, A4.3, and A4.4. The omitted categories is less than 5k for start-up capital amount. All regressions also include start quarter dummies, and those in the last two columns include six-digit NAICS industry dummies. "Demographics" includes all variables from Table A1 (except for start-up finance). The number of observations is about 55,800. Standard errors clustered by firm are in parentheses. + signifies that the coefficient at age 7 is statistically significantly different at the 5% level from the coefficient for the otherwise similar specification for age 0 (column 2 vs. 1, 4 vs. 3, 6 vs. 5, and 8 vs. 7). ^ signifies that the coefficient is statistically significantly different at the 5% level from the coefficient for the previous specification for this age (column 3 vs. 1, 4 vs. 2, 5 vs. 3, 6 vs. 4, and 8 vs. 6).

Table B2. High Growth Entrepreneurs by Gender, Age, Race, Ethnicity, and Citizenship

	No Controls		Plus Demographics		Plus Finance		Plus Industry	
Characteristic at Age 0	Top 5% Age 0	Top 5% Age 7	Top 5% Age 0	Top 5% Age 7	Top 5% Age 0	Top 5% Age 7	Top 5% Age 0	Top 5% Age 7
All female	-2.66	-2.59^{+}	-1.86$^{\wedge}$	-1.77$^{+\wedge}$	-1.17$^{\wedge}$	-1.21$^{\wedge}$	-2.25$^{\wedge}$	-2.63$^{\wedge}$
	(0.29)	(0.29)	(0.31)	(0.32)	(0.30)	(0.31)	(0.34)	(0.35)
Age 35-44	1.58	0.02^{+}	1.48$^{\wedge}$	-0.09	1.04$^{\wedge}$	-0.45$^{\wedge}$	1.31$^{\wedge}$	-0.38^{+}
	(0.31)	(0.37)	(0.31)	(0.36)	(0.30)	(0.35)	(0.30)	(0.35)
Age 45-54	2.60	-0.15	2.23$^{\wedge}$	-0.49^{+}	1.45$^{\wedge}$	-1.11$^{+\wedge}$	1.69	-1.11^{+}
	(0.33)	(0.37)	(0.33)	(0.36)	(0.33)	(0.36)	(0.33)	(0.36)
Age 55-64	3.04	-0.25	2.47$^{\wedge}$	-0.85$^{+\wedge}$	1.60$^{\wedge}$	-1.51$^{+\wedge}$	1.75	-1.54^{+}
	(0.40)	(0.42)	(0.42)	(0.43)	(0.40)	(0.42)	(0.40)	(0.42)
Age 65 or over	4.26	-0.08	3.32$^{\wedge}$	-1.02$^{+\wedge}$	1.65$^{\wedge}$	-2.28$^{+\wedge}$	1.36	-2.49^{+}
	(0.77)	(0.66)	(0.79)	(0.68)	(0.76)	(0.69)	(0.73)	(0.70)
All Hispanic	-1.61	-1.08^{+}	-0.64$^{\wedge}$	-0.33$^{\wedge}$	0.15$^{\wedge}$	0.33$^{\wedge}$	-0.94$^{\wedge}$	-0.41$^{\wedge}$
	(0.58)	(0.63)	(0.61)	(0.65)	(0.58)	(0.64)	(0.60)	(0.63)
All African American	-2.59	-1.12^{+}	-1.09$^{\wedge}$	0.21$^{+\wedge}$	-0.49$^{\wedge}$	0.71$^{\wedge}$	-1.77$^{\wedge}$	-1.25$^{\wedge}$
	(0.58)	(0.70)	(0.59)	(0.69)	(0.59)	(0.69)	(0.58)	(0.67)
All Asian	-1.17	-0.43^{+}	-0.69	-0.04	-1.18$^{\wedge}$	-0.56$^{\wedge}$	-3.69$^{\wedge}$	-1.91$^{+\wedge}$
	(0.65)	(0.69)	(0.75)	(0.78)	(0.74)	(0.77)	(0.79)	(0.82)
All other minority race	1.64	0.53	2.02	0.49	1.34$^{\wedge}$	-0.06$^{\wedge}$	0.46$^{\wedge}$	-1.02$^{\wedge}$
	(1.00)	(0.87)	(1.00)	(0.88)	(0.96)	(0.87)	(0.94)	(0.86)
All immigrant	-0.24	-0.15^{+}	0.21	0.13	-0.12$^{\wedge}$	-0.18$^{\wedge}$	-0.98$^{\wedge}$	-0.73$^{\wedge}$
	(0.41)	(0.41)	(0.50)	(0.48)	(0.49)	(0.47)	(0.49)	(0.48)

Note: Results from LPM estimation of Equation (1). Coefficients and standard errors are multiplied by 100 for ease of reading. Each regression also includes the variables in Tables A4.2, A4.3, A4.4, and A4.5. The omitted category is under 35 for owner age. All regressions also include start quarter dummies and diversity measures, "Demographics" includes all variables from Table A1 (except for start-up finance), and the last two columns include six-digit NAICS industry dummies. The number of observations is about 55,800. Standard errors clustered by firm are in parentheses. + signifies that the coefficient at age 7 is statistically significantly different at the 5% level from the coefficient for the otherwise similar specification for age 0 (column 2 vs. 1, 4 vs. 3, 6 vs. 5, and 8 vs. 7). ^ signifies that the coefficient is statistically significantly different at the 5% level from the coefficient for the previous specification for this age (column 3 vs. 1, 4 vs. 2, 5 vs. 3, 6 vs. 4, and 8 vs. 6).

Table B3. High Growth Entrepreneurs by Human Capital

	No Controls		Plus Demographics		Plus Finance		Plus Industry	
Characteristic at Age 0	Top 5% Age 0	Top 5% Age 7	Top 5% Age 0	Top 5% Age 7	Top 5% Age 0	Top 5% Age 7	Top 5% Age 0	Top 5% Age 7
Less than high school	-0.47	-1.36	-0.15^	-1.01^	0.14^	-0.75^	-0.34	-0.94
	(0.65)	(0.61)	(0.65)	(0.62)	(0.65)	(0.62)	(0.69)	(0.62)
Vocational school	-0.85	-0.65+	-0.26^	-0.16+^	-0.10	-0.03+	0.62^	0.21
	(0.53)	(0.52)	(0.52)	(0.51)	(0.51)	(0.50)	(0.52)	(0.49)
Associate degree	-0.84	-0.00+	-0.34^	0.42+^	-0.37	0.40+	-0.05	0.24
	(0.46)	(0.52)	(0.46)	(0.52)	(0.45)	(0.50)	(0.46)	(0.50)
Some college	0.27	0.60+	0.50^	0.87+^	0.27^	0.70+^	0.73^	0.83
	(0.39)	(0.40)	(0.39)	(0.39)	(0.38)	(0.39)	(0.38)	(0.39)
Bachelor's degree	2.41	2.15+	2.04	1.74+	1.19^	1.10+^	2.07^	1.29
	(0.38)	(0.38)	(0.37)	(0.37)	(0.36)	(0.36)	(0.39)	(0.38)
Graduate degree	0.31	1.36+	0.17^	1.43+^	-0.90^	0.61+^	1.48^	1.56^
	(0.38)	(0.39)	(0.38)	(0.39)	(0.37)	(0.39)	(0.47)	(0.48)
Veteran	-0.62	-1.22	-1.52^	-1.32	-1.09^	-0.99^	-1.06	-1.15
	(0.39)	(0.36)	(0.42)	(0.39)	(0.40)	(0.38)	(0.38)	(0.37)
Prior business	3.45	2.43	2.10^	1.63^	1.30^	0.96^	1.04^	0.99
	(0.25)	(0.25)	(0.25)	(0.25)	(0.24)	(0.24)	(0.24)	(0.24)

Note: Results from LPM estimation of Equation (1). Coefficients and standard errors are multiplied by 100 for ease of readingThe omitted category for owner education is high school diploma or GED. All regressions also include start quarter dummies. "Demographics" includes all variables from Table A1 (except for start-up finance), and the last two columns include six-digit NAICS industry dummies. The number of observations is about 55,800. Standard errors clustered by firm are in parentheses. + signifies that the coefficient at age 7 is statistically significantly different at the 5% level from the coefficient for the otherwise similar specification for age 0 (column 2 vs. 1, 4 vs. 3, 6 vs. 5, and 8 vs. 7). ^ signifies that the coefficient is statistically significantly different at the 5% level from the coefficient for the previous specification for this age (column 3 vs. 1, 4 vs. 2, 5 vs. 3, 6 vs. 4, and 8 vs. 6).

Table B4. High Growth Entrepreneurs by Founding Team and Diversity

	No Controls		Plus Demographics		Plus Finance		Plus Industry	
Characteristic at Age 0	Top 5% Age 0	Top 5% Age 7	Top 5% Age 0	Top 5% Age 7	Top 5% Age 0	Top 5% Age 7	Top 5% Age 0	Top 5% Age 7
Equally operated by couple	1.65	2.04+	1.19^	1.78+^	0.28^	0.94+^	-0.66^	0.15^
	(0.44)	(0.45)	(0.47)	(0.48)	(0.45)	(0.46)	(0.46)	(0.46)
Primarily husband	0.51	0.67+	0.20^	0.49+^	-0.06^	0.22^	0.14	0.28
	(0.39)	(0.40)	(0.41)	(0.42)	(0.40)	(0.41)	(0.39)	(0.40)
Primarily wife	0.76	0.99+	0.46^	0.64^	0.11^	0.23^	-1.12^	-1.34^
	(0.61)	(0.58)	(0.62)	(0.61)	(0.60)	(0.61)	(0.62)	(0.63)
Family other than couple	7.04	7.06	6.87^	7.63^	4.40^	5.52^	3.91	5.18
	(0.69)	(0.70)	(0.75)	(0.81)	(0.72)	(0.77)	(0.71)	(0.76)
Two unrelated owners	4.75	5.06+	4.56^	5.16+^	3.25^	3.98^	2.63^	3.59^
	(0.55)	(0.55)	(0.57)	(0.58)	(0.55)	(0.57)	(0.53)	(0.56)
At least 3 unrelated owners	13.08	13.62	12.37^	13.53^	7.52^	9.57^	6.45^	8.56^
	(0.96)	(0.95)	(1.00)	(1.00)	(0.98)	(1.01)	(0.96)	(1.00)
Multi-generation	5.77	5.90	1.09^	1.42^	0.22^	0.67^	-0.16	0.34
	(1.26)	(1.23)	(1.31)	(1.30)	(1.25)	(1.25)	(1.23)	(1.24)
Gender diversity	4.35	2.92	-1.70^	-3.68^	-1.01^	-3.17+^	-2.37^	-4.32^
	(0.85)	(0.77)	(0.95)	(0.89)	(0.92)	(0.88)	(0.91)	(0.87)
Ethnic diversity	0.09	1.28+	-1.86^	-1.15^	-1.32^	-0.74^	-1.41	-0.94
	(0.70)	(0.76)	(0.77)	(0.81)	(0.74)	(0.80)	(0.74)	(0.79)
Immigrant diversity	1.75	2.71	-0.20^	0.35^	-0.67^	-0.02^	-1.10	-0.44
	(0.82)	(0.87)	(0.89)	(0.91)	(0.84)	(0.90)	(0.85)	(0.89)

Note: Results from LPM estimation of Equation (1). Coefficients and standard errors are multiplied by 100 for ease of reading. All regressions also include start quarter dummies, "Demographics" includes all variables from Table A1 (except for start-up finance), and the last two columns include six-digit NAICS industry dummies. The number of observations is about 55,800. Standard errors clustered by firm are in parentheses.+ signifies that the coefficient at age 7 is statistically significantly different at the 5% level from the coefficient for the otherwise similar specification for age 0 (column 2 vs. 1, 4 vs. 3, 6 vs. 5, and 8 vs. 7). ^ signifies that the coefficient is statistically significantly different at the 5% level from the coefficient for the previous specification for this age (column 3 vs. 1, 4 vs. 2, 5 vs. 3, 6 vs. 4, and 8 vs. 6).

Appendix C. Robustness Analysis

Table C1.1 Employment Category Transition Matrices from Age 1 to Age 7 LBD 2007 Start-up Cohort: Row percent

		Age 7							
		0	1	2-4	5-9	10-19	20+	Column Total	Age 1 Share
Age 1	0	87.9	5.0	4.4	1.6	0.7	0.4	32.7	0.0
	1	58.9	25.2	12.2	2.7	0.8	0.3	23.0	5.0
	2-4	50.0	9.0	26.7	10.5	2.9	1.0	23.8	13.8
	5-9	46.2	2.8	12.8	22.9	11.5	3.7	10.8	15.2
	10-19	45.4	1.4	4.2	11.3	24.6	13.2	5.7	16.4
	20+	42.6	0.7	1.7	2.5	8.6	43.9	4.0	49.6
	Row Total	63.4	10.0	12.3	6.9	4.1	3.3	100.0	100.0

Note: Employment in the start quarter and the same quarter at age 7 are from the Business Register (BR), and firms are tracked over time using the Longitudinal Business Database (LBD). The sample is all firms starting in one of the quarters of 2007, and the sample size is about 603,000. Each cell represents the percentage of firms in the age 1 size category in the particular row that transition to the age 7 size category in the column. The Age 1 and Age 7 shares are the age 1 size category's percent of employment at age 1 and age 7, respectively.

Table C1.2 Employment Category Transition Matrices from Age 1 to Age 7 LBD 2007 Start-up Cohort: Column percent

		Age 7						
		0	1	2-4	5-9	10-19	20+	Total
Age 1	0	45.3	16.4	11.7	7.8	5.4	4.2	32.7
	1	21.3	58.0	22.8	8.9	4.2	2.1	23.0
	2-4	18.7	21.5	51.6	36.3	16.8	6.9	23.8
	5-9	7.9	3.0	11.3	36.2	30.6	12.1	10.8
	10-19	4.1	0.8	1.9	9.4	34.6	22.7	5.7
	20+	2.7	0.3	0.6	1.4	8.3	52.1	4.0
Age 7 Emp Share		0.0	2.8	9.4	12.6	15.2	60.1	100.0

Note: Employment in the start quarter and the same quarter at age 7 are from the Business Register (BR), and firms are tracked over time using the Longitudinal Business Database (LBD). The sample is all firms starting in one of the quarters of 2007, and the sample size is about 603,000. Each cell represents the percentage of firms in the age 7 size category in the particular column that have transitioned from the age 1 size category in the row.

Table C2.1 Employment Category Transition Matrices from Age 0 to Age 1: Row percent

		Age 1							
		0	1	2-4	5-9	10-19	20+	Column Total	Age 0 Share
Age 0	1	40.0	41.9	14.5	2.4	0.8	0.4	44.7	7.1
	2-4	29.6	11.1	44.7	11.5	2.2	1.0	33.4	13.9
	5-9	23.5	3.1	17.0	40.7	13.0	2.7	11.9	12.2
	10-19	21.4	3.2	4.6	16.4	42.4	12.0	6.1	12.7
	20+	22.3	0.9	2.2	2.8	12.9	59.0	4.1	54.1
	Row Total	32.7	23.0	23.8	10.8	5.7	4.0	100.0	100.0

Note: Employment in the start quarter and the same quarter at age 7 are from the Business Register (BR), and firms are tracked over time using the Longitudinal Business Database (LBD). The sample is all firms starting in one of the quarters of 2007, and the sample size is about 603,000. Each cell represents the percentage of firms in the age 0 size category in the particular row that transition to the age 1 size category in the column. The Age 0 and Age 1 shares are the age 0 size category's percent of employment at age 0 and age 1, respectively.

Table C2.2 Employment Category Transition Matrices from Age 0 to Age 1: Column percent

		Age 1						
		0	1	2-4	5-9	10-19	20+	Total
Age 0	1	54.6	81.4	27.3	10.0	6.4	4.9	44.7
	2-4	30.2	16.0	62.7	35.3	12.9	8.2	33.4
	5-9	8.5	1.6	8.5	44.5	26.9	8.1	11.9
	10-19	4.0	0.9	1.2	9.2	44.7	18.3	6.1
	20+	2.8	0.2	0.4	1.0	9.1	60.5	4.1
Age 1 Emp Share		0.0	5.0	13.8	15.2	16.4	49.6	100.0

Note: Employment in the start quarter and the same quarter at age 7 are from the Business Register (BR), and firms are tracked over time using the Longitudinal Business Database (LBD). The sample is all firms starting in one of the quarters of 2007, and the sample size is about 603,000. Each cell represents the percentage of firms in the age 1 size category in the particular column that have transitioned from the age 0 size category in the row.

Table C3.1. SBO 2007 Age 1 Estimates: High Growth Entrepreneurs by Start-up Finance

	No Controls	Plus Demographics	Plus Industry
Characteristic at Age 0	Top 5% Age 1	Top 5% Age 1	Top 5% Age 1
Capital 5k to 10k	-0.22^{+}	-0.27$^{+\wedge}$	-0.41
	(0.24)	(0.24)	(0.26)
Capital 10k to 25k	0.38^{+}	0.07$^{+\wedge}$	-0.21$^{\wedge}$
	(0.26)	(0.26)	(0.28)
Capital 25k to 50k	1.14^{+}	0.75$^{+\wedge}$	-0.03$^{\wedge}$
	(0.35)	(0.35)	(0.36)
Capital 50k to 100k	2.16^{+}	1.42$^{+\wedge}$	0.10$^{\wedge}$
	(0.39)	(0.39)	(0.41)
Capital 100k to 250k	4.87^{+}	3.63$^{\wedge}$	2.05$^{\wedge}$
	(0.52)	(0.52)	(0.52)
Capital 250k to 1m	15.16^{+}	13.21$^{\wedge}$	10.80$^{\wedge}$
	(0.95)	(0.92)	(0.87)
Capital 1m and more	32.98^{+}	29.49^{+}	26.14$^{\wedge}$
	(1.92)	(1.95)	(1.88)
Don't know amount	7.15^{+}	6.23$^{+\wedge}$	4.85$^{\wedge}$
	(0.67)	(0.66)	(0.62)
No capital needed	2.05^{+}	1.92$^{+\wedge}$	1.54$^{\wedge}$
	(0.44)	(0.44)	(0.42)

Note: Results from LPM estimation of Equation (1) at firm age 1. Coefficients and standard errors are multiplied by 100 for ease of reading. All regressions also include start quarter dummies, "Demographics" includes all variables from Table A1 (except for start-up finance), and the last two columns include six-digit NAICS industry dummies. The number of observations is about 55,800. Standard errors clustered by firm are in parentheses. + signifies that the coefficient at age 7 is statistically significantly different at the 5% level from the coefficient for the otherwise similar specification for age 0. ^ signifies that the coefficient is statistically significantly different at the 5% level from the coefficient for the previous specification for this age.

Table C3.2. SBO 2007 Age 1 Estimates: High Growth Entrepreneurs by Gender, Age, Race, Ethnicity, and Citizenship

	No Controls	Plus Demographics	Plus Finance	Plus Industry
Characteristic at Age 0	Top 5% Age 1	Top 5% Age 1	Top 5% Age 1	Top 5% Age 1
All female	-2.65^{+}	-1.87$^{+\wedge}$	-1.14$^{\wedge}$	-2.58$^{\wedge}$
	(0.29)	(0.31)	(0.30)	(0.34)
Age 35-44	1.12^{+}	1.02$^{+\wedge}$	0.57$^{\wedge}$	0.73
	(0.34)	(0.33)	(0.32)	(0.32)
Age 45-54	1.49^{+}	1.09$^{\wedge}$	0.27$^{\wedge}$	0.43^{+}
	(0.35)	(0.35)	(0.34)	(0.33)
Age 55-64	2.32^{+}	1.53	0.59$^{\wedge}$	0.65^{+}
	(0.43)	(0.44)	(0.41)	(0.41)
Age 65 or over	3.04^{+}	1.73	-0.07$^{\wedge}$	-0.26^{+}
	(0.78)	(0.79)	(0.78)	(0.76)
All Hispanic	-2.03	-0.74$^{\wedge}$	0.09$^{\wedge}$	-0.89$^{\wedge}$
	(0.57)	(0.61)	(0.59)	(0.61)
All African American	-1.57^{+}	0.02$^{+\wedge}$	0.64$^{+\wedge}$	-1.39$^{\wedge}$
	(0.71)	(0.70)	(0.70)	(0.70)
All Asian	-1.25	-0.16$^{\wedge}$	-0.62$^{\wedge}$	-2.69$^{\wedge}$
	(0.66)	(0.75)	(0.73)	(0.79)
All other minority race	0.54	1.19$^{\wedge}$	0.45$^{\wedge}$	-0.77$^{\wedge}$
	(0.89)	(0.87)	(0.87)	(0.86)
All immigrant	-0.98	-0.57	-0.91$^{\wedge}$	-1.65$^{\wedge}$
	(0.40)	(0.48)	(0.46)	(0.47)

Note: Results from LPM estimation of Equation (1) at firm age 1. Coefficients and standard errors are multiplied by 100 for ease of reading. All regressions also include start quarter dummies and diversity measures, "Demographics" includes all variables from Table A1 (except for start-up finance), and the last two columns include six-digit NAICS industry dummies. Under age 35 is the omitted age category, and all non-Hispanic white is the omitted category for race and ethnicity. The number of observations is about 55,800. Standard errors clustered by firm are in parentheses. + signifies that the coefficient at age 7 is statistically significantly different at the 5% level from the coefficient for the otherwise similar specification for age 0. ^ signifies that the coefficient is statistically significantly different at the 5% level from the coefficient for the previous specification for this age.

Table C3.3. SBO 2007 Age 1 Estimates: High Growth Entrepreneurs by Human Capital

	No Controls	Plus Demographics	Plus Finance	Plus Industry
Characteristic at Age 0	Top 5% Age 1	Top 5% Age 1	Top 5% Age 1	Top 5% Age 1
Less than high school	-0.69^{+}	-0.18$^{+\wedge}$	0.12$^{\wedge}$	-0.25
	(0.65)	(0.65)	(0.65)	(0.69)
Vocational school	-0.98^{+}	-0.44$^{+\wedge}$	-0.29	0.28$^{\wedge}$
	(0.51)	(0.51)	(0.49)	(0.48)
Some college	0.62^{+}	0.78$^{+\wedge}$	0.53$^{+\wedge}$	0.84
	(0.41)	(0.40)	(0.39)	(0.39)
Associate degree	-0.15^{+}	0.30$^{+\wedge}$	0.28^{+}	0.32
	(0.51)	(0.51)	(0.49)	(0.49)
Bachelor's degree	2.59^{+}	2.10^{+}	1.20$^{+\wedge}$	1.80$^{\wedge}$
	(0.39)	(0.38)	(0.36)	(0.38)
Graduate degree	1.02^{+}	0.84$^{+\wedge}$	-0.29$^{+\wedge}$	1.64$^{\wedge}$
	(0.39)	(0.39)	(0.38)	(0.48)
Veteran	-0.35^{+}	-1.16$^{\wedge}$	-0.69$^{\wedge}$	-0.63
	(0.44)	(0.45)	(0.43)	(0.40)
Prior business	3.46^{+}	2.27$^{+\wedge}$	1.42$^{\wedge}$	1.16$^{\wedge}$
	(0.26)	(0.25)	(0.24)	(0.23)

Note: Results from LPM estimation of Equation (1) at firm age 1. Coefficients and standard errors are multiplied by 100 for ease of reading. All regressions also include start quarter dummies, "Demographics" includes all variables from Table A1 (except for start-up finance), and the last two columns include six-digit NAICS industry dummies. High school diploma or GED is the omitted category for owner education. The number of observations is about 55,800. Standard errors clustered by firm are in parentheses. + signifies that the coefficient at age 7 is statistically significantly different at the 5% level from the coefficient for the otherwise similar specification for age 0. ^ signifies that the coefficient is statistically significantly different at the 5% level from the coefficient for the previous specification for this age.

Table C3.4. SBO 2007 Age 1 Estimates: High Growth Entrepreneurs by Founding Team

	No Controls	Plus Demographics	Plus Finance	Plus Industry
Characteristic at Age 0	Top 5% Age 1	Top 5% Age 1	Top 5% Age 1	Top 5% Age 1
Equally operated by couple	1.09^{+}	$0.67^{\wedge}$	$-0.25^{\wedge}$	$-1.21^{\wedge}$
	(0.42)	(0.46)	(0.44)	(0.44)
Primarily husband	0.52^{+}	$0.20^{+\wedge}$	$-0.07^{\wedge}$	0.03
	(0.40)	(0.42)	(0.41)	(0.40)
Primarily wife	1.01^{+}	$0.66^{\wedge}$	$0.33^{\wedge}$	$-1.28^{\wedge}$
	(0.61)	(0.63)	(0.62)	(0.64)
Family other than couple	7.04^{+}	$6.51^{\wedge}$	$3.89^{\wedge}$	$3.29^{\wedge}$
	(0.71)	(0.77)	(0.73)	(0.72)
Two unrelated owners	5.07^{+}	$4.81^{\wedge}$	$3.43^{\wedge}$	$2.83^{\wedge}$
	(0.58)	(0.59)	(0.57)	(0.54)
At least 3 unrelated owners	14.05^{+}	13.18	$7.97^{\wedge}$	$6.53^{\wedge}$
	(0.98)	(1.01)	(1.00)	(0.98)
Multi-generation	7.95^{+}	$3.36^{\wedge}$	$2.47^{\wedge}$	2.10
	(1.44)	(1.52)	(1.44)	(1.38)
Gender diversity	4.67^{+}	$-1.88^{\wedge}$	$-1.12^{\wedge}$	$-2.49^{\wedge}$
	(0.87)	(0.97)	(0.94)	(0.92)
Ethnic diversity	0.65^{+}	$-1.29^{\wedge}$	$-0.72^{\wedge}$	-0.96
	(0.76)	(0.83)	(0.80)	(0.78)
Immigrant diversity	1.71	$-0.66^{\wedge}$	$-1.17^{\wedge}$	-1.63
	(0.84)	(0.90)	(0.87)	(0.83)

Note: Results from LPM estimation of Equation (1) at firm age 1. Coefficients and standard errors are multiplied by 100 for ease of reading. All regressions also include start quarter dummies, "Demographics" includes all variables from Table A1 (except for start-up finance), and the last two columns include six-digit NAICS industry dummies. The omitted owner type category is one owner. The number of observations is about 55,800. Standard errors clustered by firm are in parentheses. $^{+}$ signifies that the coefficient at age 7 is statistically significantly different at the 5% level from the coefficient for the otherwise similar specification for age 0. ^ signifies that the coefficient is statistically significantly different at the 5% level from the coefficient for the previous specification for this age.

Table C4.1. SBO 2012 Age 0 Estimates: High Growth Entrepreneurs by Start-up Finance

	No Controls	Plus Demographics	Plus Industry
Characteristic at Age 0	Top 5% Age 0	Top 5% Age 0	Top 5% Age 0
Capital 5k to 10k	0.62^{+}	$0.69^{\wedge}$	$0.13^{\wedge}$
	(0.46)	(0.46)	(0.48)
Capital 10k to 25k	-0.05^{+}	$-0.27^{\wedge}$	$-1.31^{+\wedge}$
	(0.43)	(0.44)	(0.46)
Capital 25k to 50k	0.90^{+}	$0.62^{\wedge}$	$-0.86^{\wedge}$
	(0.55)	(0.56)	(0.60)
Capital 50k to 100k	3.02	$2.40^{\wedge}$	$0.37^{\wedge}$
	(0.73)	(0.72)	(0.76)
Capital 100k to 250k	7.65	$6.34^{\wedge}$	$2.99^{\wedge}$
	(1.04)	(1.02)	(0.98)
Capital 250k to 1m	14.86	$13.14^{\wedge}$	$9.17^{\wedge}$
	(1.68)	(1.67)	(1.59)
Capital 1m and more	17.86^{+}	$15.62^{+\wedge}$	$11.89^{+\wedge}$
	(2.87)	(2.90)	(2.90)
Don't know amount	9.24	$8.27^{\wedge}$	$6.63^{\wedge}$
	(0.88)	(0.87)	(0.81)
No capital needed	1.79^{+}	$1.41^{\wedge}$	1.09
	(0.66)	(0.65)	(0.62)

Note: Results from LPM estimation of Equation (1) using the 2012 SBO. Coefficients and standard errors are multiplied by 100 for ease of reading. All regressions also include start quarter dummies, "Demographics" includes all variables from Table A1 (except for start-up finance), and the last two columns include four-digit NAICS industry dummies. The omitted category is less than 5k for start-up capital amount. The number of observations is about 21,000. Standard errors clustered by firm are in parentheses. $^{+}$ signifies that the coefficient with the SBO 2012 cohort is statistically significantly different at the 5% level from the coefficient for an otherwise similar specification for the 2007 cohort. $^{\wedge}$ signifies that the coefficient is statistically significantly different at the 5% level from the coefficient for the previous specification.

Table C4.2. SBO 2012 Age 0 Estimates: High Growth Entrepreneurs by Gender, Age, Race, Ethnicity, and Citizenship

	No Controls	Plus Demographics	Plus Finance	Plus Industry
Characteristic at Age 0	Top 5% Age 0	Top 5% Age 0	Top 5% Age 0	Top 5% Age 0
All female	-2.94^{+}	-2.18$^{\wedge}$	-1.36$^{\wedge}$	-2.19$^{\wedge}$
	(0.42)	(0.44)	(0.43)	(0.51)
Age 35-44	1.04^{+}	0.81$^{+\wedge}$	0.44$^{+\wedge}$	0.57^{+}
	(0.51)	(0.51)	(0.50)	(0.49)
Age 45-54	2.68^{+}	2.19$^{\wedge}$	1.31$^{\wedge}$	1.31
	(0.56)	(0.55)	(0.54)	(0.53)
Age 55-64	2.21^{+}	1.27^{+}	0.36$^{+\wedge}$	0.35^{+}
	(0.62)	(0.63)	(0.62)	(0.62)
Age 65 or over	4.51	3.53	2.21$^{\wedge}$	1.18
	(1.05)	(1.08)	(1.06)	(1.02)
All Hispanic	-2.33^{+}	-0.53$^{\wedge}$	-0.09$^{\wedge}$	-0.87
	(0.65)	(0.70)	(0.70)	(0.70)
All African American	-2.56	-1.03$^{\wedge}$	-0.19$^{\wedge}$	-0.76$^{\wedge}$
	(0.78)	(0.78)	(0.77)	(0.78)
All Asian	-1.32	1.00$^{\wedge}$	0.33$^{\wedge}$	-1.86$^{\wedge}$
	(0.79)	(0.88)	(0.86)	(0.94)
All other minority race	-1.04^{+}	0.44$^{\wedge}$	0.40	0.40
	(0.78)	(0.82)	(0.82)	(0.83)
All immigrant	-2.72^{+}	-2.45^{+}	-2.47^{+}	-3.09^{+}
	(0.44)	(0.55)	(0.55)	(0.56)

Note: Results from LPM estimation of Equation (1) using the 2012 SBO. Coefficients and standard errors are multiplied by 100 for ease of reading. All regressions also include start quarter dummies, "Demographics" includes all variables from Table A1 (except for start-up finance), and the last two columns include four-digit NAICS industry dummies. All non-Hispanic white is the omitted category for race and ethnicity. The number of observations is about 21,000. Standard errors clustered by firm are in parentheses. $^{+}$ signifies that the coefficient with the SBO 2012 cohort is statistically significantly different at the 5% level from the coefficient for an otherwise similar specification for the 2007 cohort. $^{\wedge}$ signifies that the coefficient is statistically significantly different at the 5% level from the coefficient for the previous specification.

Table C4.3. SBO 2012 Age 0 Estimates: High Growth Entrepreneurs by Human Capital

	No Controls	Plus Demographics	Plus Finance	Plus Industry
Characteristic at Age 0	Top 5% Age 0	Top 5% Age 0	Top 5% Age 0	Top 5% Age 0
Less than high school	-1.90^{+}	-1.01$^{\wedge}$	-0.84	-1.70$^{\wedge}$
	(0.97)	(0.99)	(1.02)	(1.06)
Vocational school	-1.87^{+}	-1.56$^{+\wedge}$	-1.27^{+}	0.57$^{\wedge}$
	(0.71)	(0.71)	(0.71)	(0.76)
Some college	0.73^{+}	0.58	0.85$^{\wedge}$	1.08
	(0.66)	(0.65)	(0.64)	(0.65)
Associate degree	-0.43^{+}	-0.17$^{\wedge}$	0.14	0.38
	(0.85)	(0.85)	(0.83)	(0.81)
Bachelor's degree	2.32^{+}	2.01	1.56$^{\wedge}$	2.24$^{\wedge}$
	(0.62)	(0.61)	(0.60)	(0.64)
Graduate degree	0.10^{+}	0.00	-0.54$^{\wedge}$	1.63$^{\wedge}$
	(0.59)	(0.59)	(0.58)	(0.73)
Veteran	0.48	-0.91$^{\wedge}$	-0.32$^{\wedge}$	-0.34
	(0.79)	(0.82)	(0.80)	(0.78)
Prior business	4.09^{+}	2.96$^{\wedge}$	2.21$^{\wedge}$	1.95
	(0.40)	(0.40)	(0.39)	(0.38)

Note: Results from LPM estimation of Equation (1) using the 2012 SBO. Coefficients and standard errors are multiplied by 100 for ease of reading. All regressions also include start quarter dummies, "Demographics" includes all variables from Table A1 (except for start-up finance), and the last two columns include four-digit NAICS industry dummies. High school diploma or GED is the omitted category for owner education. The number of observations is about 21,000. Standard errors clustered by firm are in parentheses. $^{+}$ signifies that the coefficient with the SBO 2012 cohort is statistically significantly different at the 5% level from the coefficient for an otherwise similar specification for the 2007 cohort. $^{\wedge}$ signifies that the coefficient is statistically significantly different at the 5% level from the coefficient for the previous specification.

Table C4.4. SBO 2012 Age 0 Estimates: High Growth Entrepreneurs by Founding Team

	No Controls	Plus Demographics	Plus Finance	Plus Industry
Characteristic at Age 0	Top 5% Age 0	Top 5% Age 0	Top 5% Age 0	Top 5% Age 0
Equally operated by couple	2.52	1.28	$0.74^{\wedge}$	$-0.25^{\wedge}$
	(0.97)	(1.00)	(0.97)	(0.99)
Primarily husband	1.02	-0.41	$-0.94^{\wedge}$	-1.06
	(0.83)	(0.86)	(0.84)	(0.80)
Primarily wife	-0.95^{+}	-2.11^{+}	$-2.53^{+\wedge}$	-2.22
	(0.65)	(0.72)	(0.71)	(0.72)
Family other than couple	5.69	4.55	$3.09^{\wedge}$	1.93
	(1.13)	(1.16)	(1.13)	(1.10)
Two unrelated owners	4.77	$2.62^{\wedge}$	$1.35^{\wedge}$	0.84
	(0.93)	(1.01)	(1.00)	(0.97)
At least 3 unrelated owners	9.06^{+}	$6.13^{+\wedge}$	$3.46^{+\wedge}$	2.24^{+}
	(1.70)	(1.75)	(1.76)	(1.83)
Multi-generation	2.80^{+}	$-1.94^{\wedge}$	$-2.80^{\wedge}$	-2.94
	(1.68)	(1.75)	(1.70)	(1.73)
Gender diversity	6.72	$3.11^{+\wedge}$	3.21^{+}	2.14^{+}
	(1.53)	(1.65)	(1.62)	(1.57)
Ethnic diversity	2.75	$0.19^{\wedge}$	0.59	-0.43
	(1.37)	(1.33)	(1.31)	(1.26)
Immigrant diversity	3.53	$0.85^{\wedge}$	0.46	-0.23
	(1.53)	(1.51)	(1.48)	(1.45)

Note: Results from LPM estimation of Equation (1) using the 2012 SBO. Coefficients and standard errors are multiplied by 100 for ease of reading. All regressions also include start quarter dummies, "Demographics" includes all variables from Table A1 (except for start-up finance), and the last two columns include four-digit NAICS industry dummies. The omitted owner type category is one owner. The number of observations is about 21,000. Standard errors clustered by firm are in parentheses. $^{+}$ signifies that the coefficient with the SBO 2012 cohort is statistically significantly different at the 5% level from the coefficient for an otherwise similar specification for the 2007 cohort. $^{\wedge}$ signifies that the coefficient is statistically significantly different at the 5% level from the coefficient for the previous specification.