Age and High-Growth Entrepreneurship*

Pierre Azoulay, MIT and NBER

Benjamin F. Jones, Northwestern University and NBER

J. Daniel Kim, MIT

Javier Miranda, U.S. Census Bureau

April 2019

Abstract

Many observers, and many investors, believe that young people are especially likely to produce the most successful new firms. Integrating administrative data on firms, workers, and owners, we study startups systematically in the U.S. and find that successful entrepreneurs are middle-aged, not young. The mean age at founding for the 1-in-1,000 fastest growing new ventures is 45.0. The findings are similar when considering high-technology sectors, entrepreneurial hubs, and successful firm exits. Prior experience in the specific industry predicts much greater rates of entrepreneurial success. These findings strongly reject common hypotheses that emphasize youth as a key trait of successful entrepreneurs.

* We thank Shawn Klimek, Mark Leach, David Robinson, Scott Stern, Peter Klenow and two anonymous referees for helpful comments. We thank PCRI and Josh Lerner for access to the matched Business Register-PCRI crosswalk. Any opinions and conclusions expressed herein are those of the author(s) and do not necessarily represent the views of the U.S. Census Bureau or its staff. All results have been reviewed to ensure that no confidential information is disclosed. Contact: pazoulay@mit.edu; bjones@kellogg.northwestern.edu; jdkim@mit.edu; javier.miranda@census.gov.

"Young people are just smarter," Mark Zuckerberg, founder of Facebook

"The cutoff in investors' heads is 32…after 32, they start to be a little skeptical."
Paul Graham, venture capitalist and founder of Y Combinator[1]

I. Introduction

Entrepreneurship has long been heralded as a key driver of rising living standards (Smith 1776, Schumpeter 1942, Lucas 1978), but successful entrepreneurship is rare, with the vast majority of entrepreneurs failing to provide the major innovations or creative destruction that can drive economic growth (Glaeser 2009; Haltiwanger et al. 2013; Guzman and Stern 2017; Levine and Rubenstein 2017). In understanding entrepreneurship, and the rarity of substantial success, a key set of questions surrounds the traits of the entrepreneurs themselves. In this paper, we provide wide-ranging evidence about one trait often thought to play a central role: the founders' age.

The view that young people are especially capable of producing big ideas – whether in scientific research, invention, or entrepreneurship – is common and longstanding (see, e.g., Jones et al. 2014). Among the advantages of youth in technology and innovation, young people are sometimes argued to be cognitively sharper, less distracted by family or other responsibilities, and more capable of transformative ideas – this last in line with "Planck's Principle", whereby younger people may be less beholden to existing paradigms of thought and practice (Planck 1949; Dietrich and Srinivasan 2007, Weinberg 2006, Jones 2010, Azoulay et al. 2018). Famous individual cases such as Bill Gates, Steve Jobs, and Mark Zuckerberg show that people in their early 20s can create eventually world-leading companies. Meanwhile, venture capital firms appear to emphasize youth as a key criteria in targeting their investments, which has led to charges of "ageism" in Silicon Valley.[2] At one extreme, Peter Thiel, the co-founder of PayPal, has created a prominent fellowship

[1] Source: Nathaniel Rich, "Silicon Valley's Start-up Machine," *New York Times*, May 2, 2013.
[2] Vinod Khosla, the co-founder of Sun Microsystems and a prominent venture capitalist, has argued that "people under 35 are the people who make change happen," and "people over forty-five basically die in terms of new ideas." (source: Vivek Wadhwa, "The Case for Old Entrepreneurs," *Washington Post*, December 2, 2011). For public debate around venture capital activity and potential "ageism" see, for example "The Brutal Ageism of Tech" (Scheiber 2014).

program that provides \$100,000 grants to would-be entrepreneurs so long as they are below age 23 and drop out of school.

Despite these potential advantages, young entrepreneurs may also face substantial disadvantages. Older entrepreneurs might access greater human capital, social capital, or financial capital. Theories of entrepreneurship often take human-capital orientations (e.g., Lucas 1978; Kihlstrom and Laffont 1979; Iyigun and Owen 1998; Lazear 2004, 2005; Amaral et al. 2011), and empirical studies have found that human capital, including the acquisition of relevant market and technical knowledge, can predict entrepreneurial success (e.g., Dunn and Holtz-Eakin 2000, Fairlie and Robb 2007, Gruber et al. 2008, Chatterji 2009, Lafontaine and Shaw 2014). In deeper technological areas, young people may not have sufficient scientific knowledge to produce or manage effective R&D (e.g., Jones 2010). Age and experience may also be relevant when accessing financial capital, where younger individuals will have less time to build up capital needed to start a business and may face difficulties borrowing it (e.g., Evans and Jovanovic 1989; Stiglitz and Weiss 1981).[3] Whether such issues impose important constraints in the entrepreneurial context is less clear, especially to the extent that young entrepreneurs can overcome personal limitations by assembling effective teams, accessing third-party financing, and tapping social networks.

The empirical literature on the characteristics of highly successful entrepreneurs is limited and mixed. Various studies suggest that mean age for starting companies of all kinds (i.e., including restaurants, dry cleaners, retail shops, etc.) is in the late 30s or 40s (e.g., Dahl and Sorensen 2012, Kautonen et al. 2014), but the data in these studies are dominated by small businesses without growth ambitions and do not focus on the relatively rare start-ups with the potential to drive innovation and economic growth. Other research suggests that growth-oriented firms and the people who start them have distinct characteristics (e.g., Guzman and Stern 2017, Levine and Rubinstein 2017). Meanwhile, studies of technology firms in the U.S. find contrasting results. Roberts (1991), looking across small samples of tech entrepreneurs, finds a median founder age of 37 among 270 new ventures, while Wadhwa et al. (2008) use a telephone survey of 502 technology and engineering firms with at least \$1 million in sales and find that the mean founder

[3] In Evans and Jovanovic (1989) the entrepreneur's wealth limits the amounts of funds she can access. Empirical evidence for this mechanism continues to be debated (e.g., Holtz-Eakin et al. 1994a, 1994b; Hurst and Lusardi 2004; Andersen and Nielsen 2012; Fort et al. 2013; Adelino et al. 2015).

age was 39. Ng and Stuart (2016) connect Angel List and CrunchBase data to individual LinkedIn profiles and find, in sharp contrast, that the founding of tech ventures comes most commonly only 5 years after college graduation. Frick (2014) studies a sample of 35 VC-backed firms from the Wall Street Journal's Billion Dollar Startup Club list and finds a mean founder age of 31, echoing the popular view that the most successful and transformative new ventures come from young people (Table A1 in the online appendix further characterizes popular perceptions).

In this paper, we deploy U.S. administrative datasets to investigate the link between age and high-growth entrepreneurship in a systematic manner. By linking (a) newly available IRS K-1 data, which identifies the initial owners of pass-through firms, with (b) U.S. Census Bureau datasets regarding businesses, employees, and individuals throughout the economy as well as (c) USPTO patent databases and third-party venture-capital databases, we provide systematic new facts about founder age and entrepreneurship.

While we will include results for all new firms, our emphasis is on founders of "growth-oriented" firms that can have large economic impacts and are often associated with driving an increasing standard of living (Schumpeter 1942, Glaeser 2009). To delineate growth-oriented start-ups, we use both *ex ante* and *ex post* measures. The ex-ante measures include being a participant in a high tech sector, owning a patent, or receiving VC backing. The ex-post measures examine growth outcomes directly for each firm. Our datasets allow us to investigate multiple measures of firm growth and success at the firm level, including exceptionally high employment and sales growth, as well as exit by acquisition or initial public offering.

Our primary finding is that successful entrepreneurs are middle-aged, not young. We find no evidence to suggest that founders in their 20s are especially likely to succeed. Rather, all evidence points to founders being especially successful when starting businesses in middle age or beyond, while young founders appear disadvantaged. Across the 2.7 million founders in the U.S. between 2007-2014 who started companies that go on to hire at least one employee, the mean age for the entrepreneurs at founding is 41.9. The mean founder age for the 1 in 1,000 highest growth new ventures is 45.0. The most successful entrepreneurs in high technology sectors are of similar ages. So too are the most successful founders in entrepreneurial regions of the U.S. While the prevalence of the highest-growth companies having middle-aged founders is due in part to the prevalence of entry by the middle-aged, we further find that the "batting average" for creating successful firms is rising dramatically with age. Conditional on starting a firm, a 50-year-old founder is 1.8 times more likely to achieve upper-tail growth than a 30-year-old founder. Founders in their early 20s have the lowest likelihood of successful exit or creating a 1 in 1,000 top growth firm.

The rest of the paper is organized as follows. Section II details the newly-integrated administrative datasets that make this study possible. Section III presents our main results. Section IV presents extensions and discussion. Section V concludes.

II. Data and Measurement

Our study uses administrative data to identify the demographics of business founders in the U.S. and to track the performance of their businesses over time. Our primary datasets include administrative data from the U.S. Census Bureau's Longitudinal Business Database (LBD) and Schedule K-1 business owners data, while also integrating numerous other datasets. Detailed information about each data set is provided in the online appendix, with a summary displayed in Table A2. Below we describe how key measurement challenges can be overcome with the above databases, which enable us to analyze the demographics of business founders and track the performance of their firms over time.

Identifying New Firms. We rely on the LBD to identify startup firms. The LBD tracks both firms and their establishments over time. We follow Haltiwanger et al. (2013) and define a business's age as the age of the oldest establishment present at the first appearance of a new firm identifier. Startups are identified as de novo firms with no prior activity at any of its establishments. This approach ensures our definition of entrepreneurial firms does not include spinoffs from existing firms or new firms that are the result of the reorganization or recombination of existing businesses.[4] Note that the LBD identifies the startup year as the year when the business first hires an employee; as such the LBD startup date might differ from the legal founding date of a business. As a robustness check, we exclude businesses where the K-1 form founding date differs from the LBD age by more than two years. All results are consistent with the main findings from the full sample.

[4] We also drop age zero firms that have multiple establishments in their birth years. On average, their initial employment in year zero is unusually high relative to other new firms, suggesting that they are not de novo startups. Inspection of these startups suggest they are the result of multinational activity as well as newly created professional employer organizations.

Identifying Founders. Critical to our effort is the identification of founders. For S-corporations and partnerships, we use Form K-1 to define owners as individuals who own some portion of the firm at age zero in the LBD. We then use the W-2 data to define a founder as an owner who also works at the firm (as opposed to an investor who holds equity in the firm but does not work there). The identification of these "owner-workers" is, while traditionally very difficult in the U.S. data, straightforward in the linked administrative datasets we use.[5]

For C-corporations, we rely on two alternative approaches, as K-1 owner data is not available. For our primary analysis, we use the W-2 data to define the three highest paid workers in the first year of the firm's existence. This is the approach followed by Kerr and Kerr (2017), who argue that business owners are often among the top three initial earners in the firm.[6] Based on the S-corporation data, where ownership status can be determined with certainty, 90% of the owner-workers are in fact among the top three earners in the firm during the first year.[7] This "initial team" definition of founders can be applied to all firms. Secondarily, we will present results using the U.S. Census Annual Survey of Entrepreneurs (ASE), allowing us to look at a large subsample of C-corps for whom we can directly determine owner-workers.[8] In general, we have analyzed all of our results separately for S-Corporations (K-1 entities), partnerships (K-1 entities), and C-Corporations (non K-1 entities). Because the results are similar for each type, the main results emphasize the age findings pooled across all U.S. startups. In Section IV, we will demonstrate robustness across different ways of defining founders and different legal forms.

Identifying High-Growth Startups. We are especially interested in examining growth-oriented startups. We take two approaches. The first approach considers technology-orientation, which can

---

[5] For about 20% of new S-corporations, none of the owners work at the firm, which we interpret as businesses where the equity holders are financing a new business and running it through hired management. These firms are not included in our analysis below; we will be considering these firms more closely in further work.

[6] Kerr and Kerr (2017) use LEHD data which currently excludes Massachusetts whereas we use more comprehensive W-2 earnings records. We have separately considered our analysis using LEHD records, including different definitions of founding team based on quarterly employment data, and find very similar results as in our W-2 sample.

[7] This approach is thus good at capturing owner-workers in the sense that few are missed. However, examining the S-Corporation data, the top three earners also typically include individuals who do not have ownership stakes in the firm. Thus this "initial team" definition of founders is best thought of as a related but distinct way of capturing the important individuals in the initial life of the firm, as opposed to an exact way of capturing owner-workers. We will consider distinctions between these approaches below.

[8] The Annual Survey of Entrepreneurs (ASE) is a representative survey of U.S. businesses with paid employees and receipts of $1,000 or more.

suggest the potential for high growth. The second approach considers the actual outcome for the firm, based on the 3, 5, or 7 year time window after founding. We exclude from our analysis sole proprietors and businesses without employees.

Noting that there is no commonly accepted definition of "high tech" sectors or firms, we use three alternative definitions. First, following Hecker (2005), we define high tech sectors as industries (4-digit NAICS) with the highest share of technology-oriented workers according to the Bureau of Labor Statistics.[9] Second, we use a comprehensive match between the Census LBD and the businesses covered by the PCRI and VentureXpert databases to determine whether a given firm receives venture capital, suggesting that the firm is seen as having substantial growth potential. Third, we leverage prior research that matches the USPTO patent database with the LBD (Graham et al. forthcoming) to determine whether a firm has received a patent.

While the above measures attempt to delineate firms with substantial *potential* for growth, the LBD also allows us to quantify growth outcomes for each firm directly. Our primary outcome measures include (a) employment growth, and (b) sales growth, while we also consider (c) exit by acquisition and (d) initial public offerings. In the main text, we will emphasize employment growth, denoting a high-growth new venture as one that achieved a given threshold of employment 5 years after founding. We examine employment thresholds based on the Top 10, 5, 1, or 0.1 percentile. Analyses using sales growth are provided in the online appendix and show extremely similar results. Startups can grow and expand to become large multi-establishment corporations spanning multiple types of activities and locations. For these startups we calculate total firm employment by aggregating the establishment level records for each firm-year observation. From these firm-level measures it is straightforward to compute measures of employment growth by looking at the change in total employment over time.

Startups can also become targets for acquisition by existing firms. For example, the owner(s) of a successful venture might decide to exit by selling their idea and the assets embodied

[9] The list of Hecker (2005) includes 46 four-digit NAICS industries. An industry is considered high tech if the share of technology-oriented workers is at least twice the overall average of 4.9%. Defined by the Bureau of Labor Statistics, technology-oriented occupations are generally roles that require knowledge of science, engineering, mathematics, and/or technology typically acquired through specialized higher education.

in their firm. In this case the original firm will cease to exist as such after the acquisition.[10] Some startups will simply fail and shutdown. We separately identify acquisitions of startups by existing firms as well as shutdowns and classify these events as distinct types of firm outcomes.[11] Lastly, we use the Compustat-Business Register Bridge to identify firms that enter public equity markets through an IPO. Our measure of "successful exit" below is an indicator for acquisition or IPO ever occurring within the scope of our databases.

III. Results

We now turn to the analysis of founder age in the universe of U.S. startups delineated above. Table 1 presents the results. Focusing on the first row and first column, which shows all new ventures in the U.S., we see that the mean age at founding is 41.9. This finding is broadly consistent with other population surveys of general types of new firms. Of course, while the word "startup" may conjure the image of technology entrepreneurs in their proverbial Silicon Valley garage, the great bulk of the new ventures that constitute our universe do not match this archetype. Though our data do not include sole proprietor businesses, it is still the case that most U.S. firms do not have the ambition and/or the business model to grow and scale their business (Hurst and Pugsley, 2011).[12]

To focus on growth-oriented entrepreneurs within our universe of U.S. startups, we take several approaches. Our first set of approaches examines the nature of the startup at founding, based on technology-related criteria. Our second set of approaches examines the growth performance of the startups themselves. Given the scale of the administrative data, we can further look at intersections of these criteria to focus on narrow subgroups of firms that both grow quickly and are in high-technology areas.

III.A Ex-Ante Growth-Orientation

---

[10] In the LBD these firms' establishments will take on the acquiring firms' identifiers.
[11] To distinguish successful acquisitions (i.e., those that generate positive returns for investors) from fire sale acquisitions, we drop observations for which total employment after the acquisition is lower than initial employment.
[12] While excluded from the analysis, our data show that the average age of new sole proprietors in 2010 was 44.8, significantly older than the rest of the population.

The results for different measures of growth-orientation are found in columns (2)-(4) of Table 1. We see that focusing on "high-tech" does not substantively affect mean founder age compared to the overall U.S. sample. Depending on the definition of high-technology, mean founder age now ranges from 41.9 to 44.6, with founders in high-tech sectors (43.2) and founders of patenting firms (44.6) appearing somewhat older on average than founders in the U.S. overall.

We can further partition the data geographically and consider California, Massachusetts, and New York separately given that these three states account for significant portions of high-growth startup activity in the U.S. (see Chen et al. 2010 with respect to VC-backed startups). In addition, we can examine regions with the most entrepreneurial activity at the zip code level. Using the Entrepreneurial Quality Index developed by Guzman and Stern (2017), we define entrepreneurial hubs as the 50 zip codes with the highest entrepreneurial quality. We also look specifically at Silicon Valley, considering all new ventures in the zip codes of Santa Clara and San Mateo counties.

Taking the overall population of new ventures (column 1), we see little variation with geography. Even when looking at the zip codes with the most growth-oriented new ventures, the mean founder age is 40.8, or approximately 1 year younger than the U.S. population average. One interpretation of this result may be that, even in entrepreneurial regions, most new firms are not in technology or growth-oriented sectors. However, reading across columns and rows in the table, we can further examine the intersection of geography with technology or growth-orientation. Remarkably, we see only modest differences in age. Mean founder ages rarely dip much below age 40, let alone ages 35, 30, or 25. The only category where the mean ages appear (modestly) below age 40 is when the firm has VC-backing. The youngest category is VC-backed firms in New York, where the mean founder age was 38.7. More generally, across the various narrow cuts in Table 2, the mean age ranges from 38.7 to 45.3. Put another way, even when reducing the set of 2.7 million founders to the 1,900 associated with firms that are both in entrepreneurial hubs and receive VC backing, the mean age at founding is 39.5. Meanwhile, founders in high-tech employment sectors tend to be slightly older than the U.S.-wide average, and founders of patenting firms are the oldest of all, with an average age of 44.3 in Silicon Valley and 43.8 in the entrepreneurial hubs.

III.B Ex-Post High-Performance Firms

It may still be that younger founders produce the highest performance new firms. Our second approach considers firm-level outcomes. The capacity to examine firm performance draws on the strengths of the LBD, which provides employees and sales for each firm, as well as indicating exit by acquisition and, via the Compustat Bridge, initial public offerings. A potential limitation in the intersection of our databases is that we have a limited time-period in which we can examine firm performance. Here we will focus on growth outcomes five years after the hiring of the first employee.[13]

To delineate "successful" entrepreneurs within the population of new ventures, we focus on the upper tail of the new ventures' employment growth. Specifically, we examine firms alternatively in the Top 10%, Top 5%, Top 1%, and Top 0.1% of growth. We complement these employment-based growth measures with a metric tracking whether these ventures ever exited by acquisition or IPO within our sample period.

Table 2 presents founder age across a range of upper-tail performance definitions. We see that more successful startups have, if anything, slightly older founders on average. For example, the 1,700 founders of the fastest growing new ventures (the top 0.1%) in our universe of U.S. firms had an average age at founding of 45.0 (compared to 43.7 for the top 1% and 42.1 for the top 5%). Regardless of the measure of technology-intensiveness chosen, we see older founders as we move toward upper-tail performance, especially for the top 1 in 100 or top 1 in 1,000 firms, as well as for founders with successful exits. This evidence is at odds with the conventional wisdom that successful founders skew younger.

III.C Founder Age Distributions

One limitation of the foregoing results is that they only shed light on mean founder age. While mean age provides a standard summary statistic, and one that we can compare across technology-intensity, regions, and outcome measures, investigating the entire age distribution may reveal bands of age where founder activity is especially intense or founders are especially successful.

---

[13] Using 3-year windows and 7-year windows shows broadly similar results.

Figure 1 presents the full founder age distributions, for the founders of all U.S. firms (blue line) and for Top 1% firms by employee growth after five years (red line).[14] Studying all founders, the age distribution is single peaked, with a relatively flat plateau at ages 37-43. Studying founders of high-growth firms, the founder age distribution shifts systematically to the right. Thus, the highest-growth new firms not only appear to come from those in middle-age, but also tend to come at even older ages than the background age distribution for founders would imply. Prior to the late 30s, the frequency of successful founders is well below the frequency of these founders in the population. Starting in the late 30s, and especially by the mid-to-late 40s, the frequency of successful founders is substantially greater than the frequency of these founders in the population. A similar peak in middle age appears when comparing the founder age distribution against the underlying workforce age distribution as opposed to the population as a whole (Figure A1).

III.D The Likelihood of Success

Our previous results have demonstrated that growth-oriented start-up founders in the US economy tend to be middle-aged, not young. Thus, when asking where most high-growth or technology-intensive firms in the U.S. come from, the answer is "middle aged people." However, an equally important question is to ask how the probability of entrepreneurial success changes with founder age, conditional on starting a new firm. This statistic may be more informative for an individual considering founding a company or for investors deciding where to place their bets. For example, if two founders (of two distinct firms) come to pitch their idea to a venture capitalist, and all the venture capitalist knows is these founders' ages, which founder would be more likely to produce an upper-tail growth outcome?

To examine the relationship between the likelihood of success and age, we run linear probability models where an indicator for "success" is regressed on a full set of founder age fixed effects (age 20 and below is the omitted category). We graph each age coefficient and the associated 95% confidence interval in Figure 2.[15] Our success indicators are (a) exit by acquisition or IPO and (b) employment in the top 0.1% measured here at 5 years from founding.

[14] Appendix Figure A4 presents analyses using upper tail sales growth instead of employment growth and shows similar results.
[15] The regressions calculate robust standard errors, clustered at the new venture level.

Figure 2A considers successful exits, which occurs for roughly 4,000 (or 0.15%) of the founders in our universe. We see that the relationship between age and successful exit is monotonically increasing up until about age 60 and declining slightly thereafter. A founder at age 50 is approximately twice as likely to experience a successful exit compared to a founder at age 30. Figure 2B replicates this analysis using Top 0.1% employment growth as the success metric. Here again, success probabilities are increasing with age, though the individual age coefficients are estimated less precisely. Similar to the exit results, a founder at age 50 is approximately twice as likely to achieve upper-tail employment growth compared to a founder at age 30.[16]

Overall, we see that younger founders appear strongly *disadvantaged* in their tendency to produce the highest-growth companies. That said, there is a hint of some interesting age thresholds and plateaus in the data. Below age 25, founders appear to do badly (or rather, do well extremely rarely), but there is a sharp increase in performance at age 25. Between ages 25 and 35, performance seems fairly flat. However, starting after age 35 we see increased success probabilities, now outpacing the 25-year-olds. Another large surge in performance comes at age 46 and is sustained toward age 60.

IV. Extensions and Discussion

In this section, we provide secondary results and discussion to further characterize and help interpret the main findings of Section III.

IV.A Robustness across Sectors, Founding Year, Legal Form, and Founding Team Definition

The data can be cut several additional ways to further establish robustness of the main results. First, we explore heterogeneity across industries. Table A3 documents some substantial differences across sectors in the mean age of founders. Yet there is no sector, including in computing, where the mean founder ages are below 38, and only 3 of the 315 NAICS-4 digit sectors show a mean founder age below 40. Second, we consider founder age by calendar year, in part to see if the findings are robust outside the Great Recession, which occurs in our sample period. Using ex-ante or ex-post growth orientation, we find similar age results looking at calendar years individually from 2007-2014 (Table A4). Third, we disaggregate the results by legal form

[16] Results (not shown) controlling for industry are virtually unchanged.

and across definitions of the founding team (Figure A2). We see that the highest-growth firms are started by individuals in middle age and beyond regardless of legal form or founding team definition.

IV.B Age Differences within Founding Teams

We further examine age variation within founding teams. To the extent that different members of a founding team play different roles, it is theoretically possible that the youngest members play outsized roles. Further, successful firms might feature founding teams with heterogeneous ages, possibly leveraging advantages of both youth and experience. However, looking at the *youngest* member of successful founding teams, a pre-middle-age tendency does not emerge (Table A5). For the Top 0.1% of new ventures, the youngest members center in the late 30s and early 40s (while the oldest members center in the late 40s and early 50s).

IV.C Entrepreneurial Outliers

Although we have looked at the Top 0.1% of firms and the rare outcome of successful acquisition or IPO, one might still wonder if even more extreme upper-tail outliers are the province of the very young. More precisely, several cases of extreme entrepreneurial success in the software and IT sectors have prominently featured very young founders (e.g., Steve Jobs, Bill Gates, and Mark Zuckerberg). One response to this observation is to balance the ledger by noting cases of extraordinary successes featuring older founders. For example, Herbert Boyer was age 40 when, based on his genetic engineering breakthroughs, he founded Genentech (which would eventually be acquired for $47 billion), and David Duffield was 64 when he founded Workday (which currently has a market capitalization of $43 billion).

At the same time, a subtler but perhaps more important response may lie among the greatest young founders themselves. Namely, the claim that young people are especially good at starting companies is a *within person* claim. That is, a given individual is thought to be "better" when s/he is young (e.g., when s/he may have greater energy, deductive abilities, originality, etc.). If so, then we would expect great young entrepreneurs to become "worse" when they age. At a cursory level, this seems doubtful. Elon Musk's Tesla and SpaceX seem no less visionary than his earlier ventures, Zip2 and X.com. Steve Jobs and Apple computer appeared to find their blockbuster innovation with the iPhone, introduced when Jobs was 52. Jeff Bezos and Amazon have moved far beyond selling books online. These examples suggest that these prominent founders themselves may not have peaked when very young.

To examine this idea quantitatively, we studied the forward 5-year stock price multiple as a function of founder age for each of Microsoft, Apple, Amazon, and Google.[17] This analysis allows us to examine whether the additional growth in market valuation tends to decline as these individuals age. We see no such tendency (Figure A3). In fact, the five-year multiples tend to rise toward middle age. The peaks come at age 48 (Steve Jobs), age 39 (Bill Gates), age 45 (Jeff Bezos), and age 36 (Sergei Brin and Larry Page).

Because many forces influence the stock prices of firms, interpreting these results requires substantial caution. With this important caveat in mind, however, the patterns may suggest a potential reconciliation between the existence of great young entrepreneurs and the advantages of middle age. Namely, extremely talented people may also be extremely talented when young. These individuals may succeed at very young ages, even when people (including these young successes) get better with age. Thus there is no fundamental tension between the existence of great young entrepreneurs and a general tendency for founders to reach their peak entrepreneurial potential later in life.

IV.D Industry Experience

Among successful entrepreneurs more broadly, we further consider the idea that capabilities may increase with experience by consulting prior employment histories. Using the LEHD to link 2.5 million founders to their prior work experience, we examine, for every founder, whether the individual has prior work experience in the specific sector of the start-up. Overall, the results (Table 3) indicate that founders with both closer and longer experience in the specific industrial sector of the start-up see substantially greater success rates. For achieving a 1 in 1,000 highest-growth firm, having no experience in the 2-digit level industry leads to a success rate of 0.11%, while having at least three years of experience in the start-up's industry shows success rates rising to 0.22% (NAICS2 experience), 0.24% (NAICS4 experience), and

[17] The stock price multiple is the ratio of the closing stock price five years in the future to the January 1st closing stock price in the current year. The stock price series are post IPO and account for dividends and splits. While Facebook would be a natural addition to this quartet of firms, the stock price series is too short as yet to allow such analysis.

0.26% (NAICS6 experience). These findings are the opposite of stories that emphasize an outsider advantage for founders – which is a primary rationale underlying the broader belief that young people will produce the highest-growth firms.

IV.E Prior Wages

We can further incorporate individuals' prior W-2 wages into the decision to start new firms. Net of wage controls, we find that entry still peaks in middle age (Figure A5). At the same time, wages positively predict success. Individuals who start the highest-growth firms typically have very high prior wages (Figure A6), so that these individuals have outsized success both in the labor market and in founding firms. This finding is consistent with upper-tail founders having high skill; it is also consistent with the idea that high-growth founders set a high bar for entry into entrepreneurship, given a high opportunity cost of leaving the ordinary labor market behind.

IV.F Venture Capital Behavior

We also see that venture capitalists tend to bet on relatively young founders. Given that younger founders have substantially lower batting averages (e.g., see Figure 2), the founder-age tendency in VC investments may be surprising. VCs may thus be seen as making bad bets, which may be consistent with empirical findings suggesting that VCs have trouble predicting success and have earned low returns (Kaplan and Lerner 2010, Kerr et al. 2014). However, young founders may also be more in need of early-stage external finance, thus leading to this relationship. More subtly, and noting that VCs are seeking high returns, which is not identical to high growth, it may be that younger founders tend to sell their equity at lower prices, and thus VCs are making optimal return decisions. Teasing apart why VCs bet young is an interesting area for further work. We can say now however that venture capital, a major source of early-stage financing that can help drive creative destruction and economy-wide growth, does not currently appear allocated to the firms with the highest growth potential.

V. Conclusion

Researchers, policymakers, investors, and entrepreneurs themselves all strive to understand entrepreneurial traits that predict the creation of successful new firms. This paper has focused on founder age, which is often thought to be a key predictor of entrepreneurial success. We find that age indeed predicts success, and sharply, but in the opposite way that many propose. The highest success rates in entrepreneurship come from founders in middle age and beyond.

These findings are consistent with theories in which key entrepreneurial resources (such as human capital, financial capital, and social capital) accumulate with age. Mechanisms by which young people are proposed to have advantages (such as energy or originality) may still be operating, but if so they appear to be overwhelmed by other forces. Future work can explore how variation in specific founder traits predict entrepreneurial entry and success, further informing underlying theories for the life cycle of entrepreneurs and provide additional capacity to predict entrepreneurial success. More broadly, new administrative datasets linking founder traits and business outcomes promise to further reveal core facts about the high-growth new ventures that can drive economic growth and the advance of socioeconomic prosperity.

References

Adelino, Manuel and Antoinette Schoar and Felipe Severino (2015), “House prices, collateral, and self-employment.” Journal of Financial Economics, Volume 117, Issue 2, Pages 288–306.

Amaral, A. Miguel, Rui Baptista and Francisco Lima (2011), “Serial entrepreneurship: Impact of human capital on time to re-entry.” Small Business Economics, 37, 1-21.

Andersen, Steffen and Kasper Meisner Nielsen (2012), "Ability or Finances as Constraints on Entrepreneurship? Evidence from Survival Rates in a Natural Experiment." The Review of Financial Studies, Volume 25, Issue 12, 1 December, Pages 3684–3710.

Azoulay, P., Fons-Rosen, C., & Zivin, J. S. G. (2018). Does science advance one funeral at a time? (No. w21788). National Bureau of Economic Research.

Chatterji , Aaron K (2009) “Spawned with a Silver Spoon? Entrepreneurial Performance and Innovation in the Medical Device Industry.” Strategic Management Journal, Vol. 30, No. 2 (Feb., 2009), pp. 185-206.

Chen, Henry, Paul A. Gompers, Anna Kovner, and Josh Lerner (2010), "Buy Local? The Geography of Successful Venture Capital Expansion." Journal of Urban Economics 67:1. .

Dahl, M. S., & Sorenson, O. (2012). Home sweet home: Entrepreneurs' location choices and the performance of their ventures. Management science, 58(6), 1059-1071.

Dietrich, Arne. and Narayanan Srinivasan (2007), "The Optimal Age to Start a Revolution." Journal of Creative Behavior, 41:1, 54-74.

Dunn, Thomas A, and Douglas J. Holtz-Eakin. 2000. "Financial Capital, Human Capital, and the Transition to Self-Employment: Evidence from Intergenerational Links." Journal of Labor Economics, Vol. 18, No. 2, pp. 282-305.

Evans, David S. and Boyan Jovanovic (1989), “An Estimated Model of Entrepreneurial Choice under Liquidity Constraints.” Journal of Political Economy, Vol. 97, No. 4 pp. 808-827.

Fairlie, Robert W., and Alicia M. Robb. 2007. "Families, Human Capital, and Small Business: Evidence from the Characteristics of Business Owners Survey," (with Alicia Robb), Industrial and Labor Relations Review, 60(2): 225-245.

Fort, Teresa C., John C. Haltiwanger, Ron S. Jarmin and Javier Miranda (2013), “How firms respond to business cycles: The role of firm age and firm size.” IMF Economic Review 61 (3), 520-559.

Foster, Lucia and Patrice Norman. “The Annual Survey of Entrepreneurs: An Update.” CES working paper 17-46, June 2017.

Frick, Walter (2014) “How Old Are Silicon Valley’s Top Founders? Here’s the Data.” Harvard Business Review.

Glaeser, Edward L. (2009), “Entrepreneurship and the City”. In *Entrepreneurship and Openness: Theory and Evidence*, edited by David B. Audretsch, Robert Litan, Robert J. Strom.

Graham, Stuart J. H., Cheryl Grim, Tariqul Islam, Alan C. Marco and Javier Miranda (forthcoming), “Business Dynamics of Innovating Firms: Linking U.S. Patents with Administrative Data on Workers and Firms.” Journal of Economics and Management Strategy.

Gruber, Marc, Ian MacMillan and James Thompson (2008) “Look Before You Leap: Market Opportunity Identification in Emerging Technology Firms.” Management Science Vol. 54, No. 9, September 2008, pp. 1652-1665.

Guzman, Jorge and Scott Stern (2017), “Nowcasting and placecasting entrepreneurial quality and performance” Chapter in NBER book *Measuring Entrepreneurial Businesses: Current Knowledge and Challenges*, pages 11-62 National Bureau of Economic Research, Inc., John Haltiwanger, Erik Hurst, Javier Miranda, and Antoinette Schoar, editors.

Haltiwanger, J., Jarmin, R. and Miranda, J. (2013), "Who Creates Jobs? Small vs. Large vs. Young." Review of Economics and Statistics, 95:2, 347-361.

Hecker, D. E. (2005), "High-Technology Employment: a NAICS-Based Update," Monthly Lab. Rev., 128, 57.

Holtz-Eakin, Douglas and David Joulfaian and Harvey Rosen (1994a), "Entrepreneurial Decisions and Liquidity Constraints." RAND Journal of Economics, vol. 25, issue 2, 334-347.

Holtz-Eakin, Douglas and David Joulfaian and Harvey Rosen (1994b), "Sticking It Out: Entrepreneurial Survival and Liquidity Constraints." Journal of Political Economy, 1994, vol. 102, issue 1, 53-75.

Hurst, Erik, and Benjamin Wild Pugsley (2011) “What Do Small Businesses Do?” Brookings Papers on Economic Activity, no. 2, pp. 73–142.

Hurst, Erik, and Annamaria Lusardi (2004), “Liquidity Constraints, Household Wealth, and Entrepreneurship.” Journal of Political Economy, vol. 112, no. 2, pp. 319–347.

Iyigun, Murat F. and Ann L. Owen “Risk, Entrepreneurship, and Human-Capital Accumulation.” The American Economic Review, Vol. 88, No. 2, Papers and Proceedings of the Hundred and Tenth Annual Meeting of the American Economic Association (May, 1998), pp. 454-457.

Jones, Benjamin F (2010), “Age and Great Invention” The Review of Economics and Statistics 2010 92:1, 1-14.

Jones, Benjamin F, Reedy, EJ & Weinberg, BA (2014), Age and Scientific Genius. in DK Simonton (ed.), The Wiley Handbook of Genius. John Wiley & Sons, Inc., pp. 422-450.

Kaplan, Steven N. and Josh Lerner (2010), "It Ain't Broke: The Past, Present, and Future of Venture Capital." *Journal of Applied Corporate Finance*, 22(2), pp. 36-47.

Kautonen, T., Down, S. & Minniti, M. (2014), "Ageing and entrepreneurial preferences." Small Bus Econ (2014) 42: 579.

Kerr, William R., Ramana Nanda, and Matthew Rhodes-Kropf (2014), "Entrepreneurship as Experimentation." Journal of Economic Perspectives—Volume 28, Number, 25–48.

Kerr, Sari P. and William R. Kerr (2017), "Immigrant Entrepreneurship." Chapter in forthcoming NBER book Measuring Entrepreneurial Businesses: Current Knowledge and Challenges, John Haltiwanger, Erik Hurst, Javier Miranda, and Antoinette Schoar, editors.

Kihlstrom, Richard E. and Jean-Jacques Laffont (1979), "A General Equilibrium Entrepreneurial Theory of Firm Formation Based on Risk Aversion." Journal of Political Economy vol 87:4, 719-748.

Lafontaine, Francine and Kathryn Shaw (2014) "Serial Entrepreneurship: Learning by Doing?" NBER Working Paper No. 20312.

Lazear, E. P. (2004) "Balanced skills and entrepreneurship." American Economic Review, Papers & Proceedings, 94, 208-11.

Lazear, Edward P. (2005) "Entrepreneurship." Journal of Labor Economics, 2005, vol. 23, no. 4

Levine, Ross, & Rubinstein, Yona (2017). Smart and illicit: who becomes an entrepreneur and do they earn more?. The Quarterly Journal of Economics, 132(2), 963-1018.

Lucas, Robert E. "On the Size Distribution of Business Firms." The Bell Journal of Economics, vol. 9, no. 2, 1978, pp. 508–523.

Ng, Weiyi and Toby E. Stuart (2016), "Of Hobos and Highflyers: Disentangling the Classes and Careers of Technology-based Entrepreneurs." Working paper, UC Berkeley.

Planck, Max (1949) "The Meaning and Limits of Exact Science." Science 30 Sep 1949: Vol. 110, Issue 2857, pp. 319-327.

Rich, Nathaniel (2013), "Silicon Valley's Start-up Machine," *New York Times*, May 2.

Roberts, E. B. (1991). Entrepreneurs in high technology: Lessons from MIT and beyond. Oxford University Press.

Scheiber, Noam (2014), “The Brutal Ageism of Tech: Years of experience, plenty of talent, completely obsolete.” The New Republic, March 23, 2014.

Schumpeter, Joseph A. (1942) "Creative Destruction" From Capitalism, Socialism and Democracy (New York: Harper, 1975) [orig. pub. 1942]

Smith, Adam (1776), *The Wealth of Nations*. W. Strahan and T. Cadell, London.

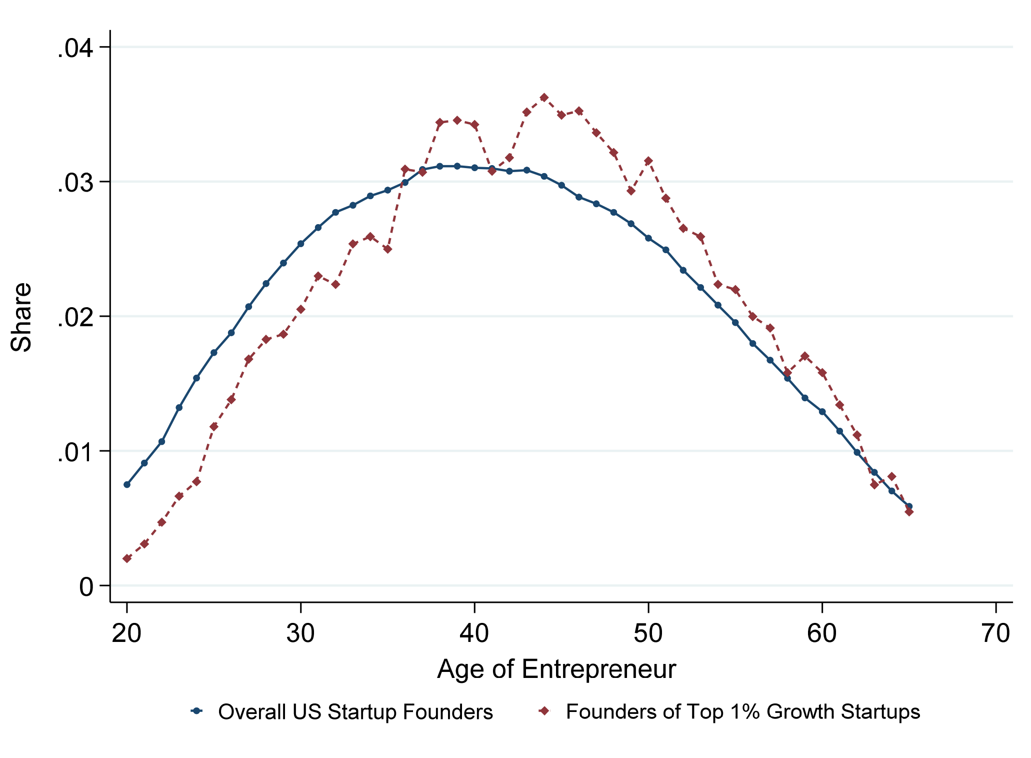
Stiglitz, Joseph E., and Andrew Weiss (1981), “Credit Rationing in Markets with Imperfect Information.” The American Economic Review, vol. 71, no. 3, pp. 393–410.

Wadhwa, Vivek (2011), “Over the Hill at 40.” ASEE Prism 21 (1), 32-32.

Wadhwa, Vivek , Richard Freeman and Ben Rissing (2008), “Education and Tech Entrepreneurship.” Kauffman Foundation Brief.

Weinberg, Bruce A. (2006). “Which labor economists invested in human capital? Geography, vintage, and participation in scientific revolutions.” Working Paper, Ohio State University mimeo.

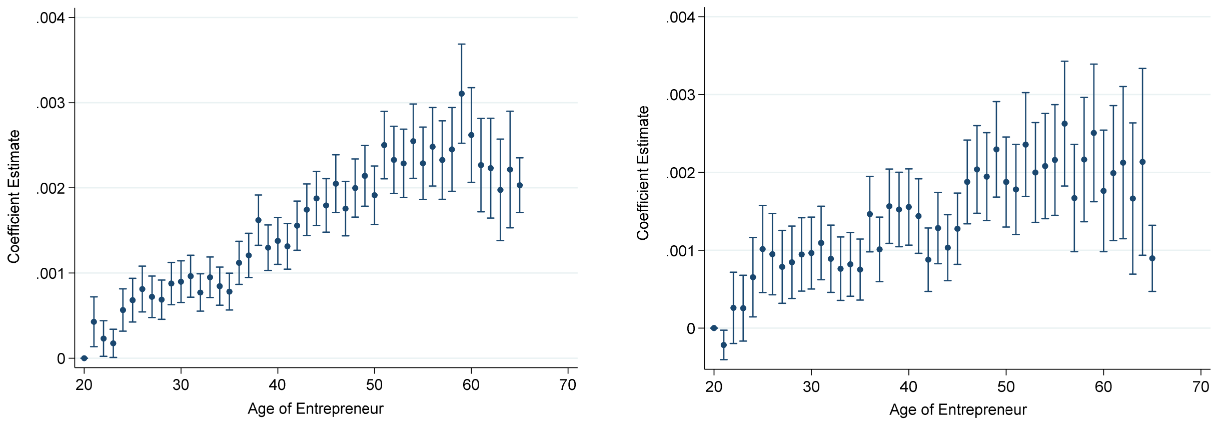
Figure 1: Founder Age Distribution: All Startups and High Growth Startups



Source: Authors calculations based on W-2 earnings records, form K-1 and Longitudinal Business Database.
Notes: This set of kernel density plots shows the age distribution of startup founders (at year of founding) in the US. Each bin represents an age cohort. Ages between 20 and 65 are incorporated in the plots. The blue (left) plot incorporates all founders of new C-corporations, S-corporations, and Partnerships with employees founded between 2007 and 2014 as identified in the Longitudinal Business Database (LBD). The red (right) plot represents founders of the top 1% growth firms founded over the 2007-2009 period. The top 1% employment growth threshold value is calculated for each yearly cohort based on the raw employment figures from the LBD in the five years after the birth of the firm.

Figure 2: Likelihood of Extreme Success, Conditional on Starting a Firm

Fig. 2A: Probability of Successful Exit (IPO or acquisition), by Age *2B: Probability of Top 0.1% Employment at 5 Years, by Age*



Source: Authors calculations based on W-2 earnings records, form K-1, Longitudinal Business Database and Compustat for firms founded over the 2007-2009 period.
Notes: OLS regression coefficients from estimating the likelihood of extreme firm success on a series of age indicators are shown. Ages 19 and below are grouped as 19 while ages 66 and above at grouped as 66. IPO data are sourced from Compustat. Acquisitions are based on firm ownership changes in the Longitudinal Business Database (LBD). Top 0.1% employment outcomes are calculated based on five-year employment growth in the LBD.

Table 1: Founder Age – Averages across U.S. and by Technology Definition

	All Startups	High Tech Employment	VC-Backed Firms	Patenting Firms
US (entire)	41.9	43.2	41.9	44.6
	(12)	(11.5)	(10.6)	(11.3)
	2,658,000	334,000	11,000	10,000
California	41.7	42.1	39.6	43.9
	(12)	(11.3)	(10)	(11)
	374,000	61,700	4,000	3,000
Massachusetts	41.7	43.2	42.3	45.3
	(11.8)	(11.2)	(9.8)	(10.6)
	52,000	8,100	900	400
New York	41.4	41.8	38.7	42.7
	(11.6)	(11.6)	(10.1)	(11.4)
	276,000	22,600	800	600
Silicon Valley	41.6	41.5	40.2	44.3
	(11.4)	(10.3)	(9.7)	(9.8)
	32,000	11,700	1,700	900
Entrepreneurial hubs	40.8	40.5	39.5	43.8
	(11.3)	(10.6)	(9.8)	(10.2)
	23,000	9,300	1,900	700

Notes: Mean founder age is shown in the first row, standard deviation in parentheses, and observation count in the third row. Data incorporates all C-corporations, S-corporations, and Partnerships founded over 2007-2014. Based on the Longitudinal Business Database (LBD), only new firms from each year are included. High tech sectors in column 2 are defined at the 4-digit NAICS level (see text). Column 3 represents firms that ever receive venture capital. Column 4 represents firms that are ever granted a patent, which is derived from the Longitudinal Linked Patent-Business Database. Silicon Valley is defined as zip codes in Santa Clara and San Mateo counties. Entrepreneurial hubs are defined as zip codes with the highest entrepreneurial quality as defined by Guzman and Stern (2017). Counts are rounded to comply with disclosure rules.

Table 2: Founder Age and Success — Upper Tail Growth or Acquisition

	All Startups	Top 10%	Top 5%	Top 1%	Top 0.1%	Successfully Exited Startups
US (entire)	41.8	41.6	42.1	43.7	45.0	46.7
	(11.9)	(11.5)	(11.5)	(11.1)	(10.7)	(10.6)
	1,079,000	126,000	62,000	13,000	1,700	4,000
Tech Employment	43.2	42.1	42.3	43.6	45.9	48.4
	(11.3)	(10.5)	(10.5)	(10)	(9.6)	(9.8)
	132,000	13,000	7,800	2,200	400	1,100
VC-Backed Firms	42.4	42.3	42.5	43.3	43.4	47.9
	(10.3)	(10.1)	(10.1)	(10)	(10.1)	(9.5)
	6,600	2,500	2,000	800	140	180
Patenting Firms	44.4	44.4	44.6	45.0	46.2	49.3
	(11.1)	(10.4)	(9.9)	(9.2)	(9.7)	(10.1)
	7,000	1,900	1,300	500	90	200

Notes: Mean founder age is shown in the first row, standard deviation in parentheses, and observation count in the third row. Data incorporates all C-corporations, S-corporations, and Partnerships founded over 2007-2009 in the Longitudinal Business Database (LBD), for which we can observe 5 years of performance data after founding. Only new firms from each year are included. Employment growth is measured using the 5-year window. Tech Employment consists of NAICS-4 sectors with high shares of STEM-trained workers. Counts are rounded to comply with disclosure rules.

Table 3: Industry-Specific Experience and Growth Outcomes

Panel A: Founders with Work Experience in Startup's 2-Digit Industry Classification

	Top 10%	Top 5%	Top 1%	Top 0.1%	Successful Exit
NAICS-2 Experience					
Never	8.6%	4.1%	0.9%	0.11%	0.13%
1-2 years	10.1%	4.8%	1.0%	0.11%	0.10%
>= 3 years	15.0%	7.7%	1.7%	0.22%	0.20%

Panel B: Founders with Work Experience in Startup's 4-Digit Industry Classification

	Top 10%	Top 5%	Top 1%	Top 0.1%	Successful Exit
NAICS-4 Experience					
Never	9.1%	4.5%	1.0%	0.12%	0.14%
1-2 years	11.6%	5.6%	1.1%	0.14%	0.12%
>= 3 years	16.8%	8.5%	1.7%	0.24%	0.20%

Panel C: Founders with Work Experience in Startup's 6-Digit Industry Classification

	Top 10%	Top 5%	Top 1%	Top 0.1%	Successful Exit
NAICS-6 Experience					
Never	9.4%	4.6%	1.0%	0.12%	0.13%
1-2 years	12.6%	6.0%	1.2%	0.15%	0.13%
>= 3 years	17.7%	9.0%	1.8%	0.26%	0.21%

Notes: Data incorporates all C-corporations, S-corporations, and Partnerships founded over 2007-2009 in the Longitudinal Business Database (LBD), for which we can observe 5 years of performance data after founding. Growth outcomes are determined by employment growth, using the 5-year window after founding.

Supplementary Online Appendices

Appendix I: Further Description of Data Sets and Methods

This appendix provides additional details regarding the datasets used in this paper. Table A2 provides a summary of the datasets and their key variables. Many data sets are available to researchers through Census approved projects and accessible through Federal Statistical Research Data Centers (FSRDC), as further indicated in the table. The Schedule K-1 and Form W-2 datasets are currently accessible only by U.S. Census employees who have been granted access through approved internal projects.

The Longitudinal Business Database (LBD)

The LBD is an establishment-level longitudinal database tracking all establishments and firms in the US with at least one employee. Starting in 1976 and updated annually, the LBD currently covers years through 2015. The LBD is sourced from administrative income and payroll filings and enhanced with Census collections, including the Economic Census and the Company Organization Survey.

Key variables in the LBD include payroll, employment, industry, location (including state and county), and legal form of organization. Establishment and firm identifiers allow us to aggregate establishment-level information to the firm to identify firm-level employment and payroll. The ability to track establishments over time makes it possible to identify de novo firms (startups) distinctly from firms that emerge from corporate restructuring or M&A activity. Startups are defined as single-unit firms during the year in which the firm first appears in the LBD with at least 1 employee. In order to identify M&A activity using the LBD, we manually track changes in firm ownership. More specifically, we flag a firm ownership change when all of the existing establishments in a firm simultaneously receive a new firm identifier in the following year. In order to ascertain that the firm ownership changes are the result M&A rather than corporate expansions, we impose the following conditions: (1) the owning firm is an incumbent firm that exists in the LBD prior to the ownership change; (2) the original EIN and names of the establishments prior to the acquisitions differ from those of the new owner's prior to the acquisition. For additional information regarding the LBD, see Jarmin and Miranda (2002).

Schedule K-1 (Form 1065/1120)

Schedule K-1 is a tax form used to report business income or loss for owners of S-corporations and partnerships. Partnerships and S-corporations are pass-through entities, meaning that their profits are not taxed at the entity level but rather as they flow through to the owners. The Schedule K-1 reports the amount of income passed through to each party. Partnerships and S-corporations file a separate K-1 form for each of their owners and are required to account for 100% of profits. The availability of both employer identification numbers (EIN) and person identification numbers (SSN) allow us to identify all the owners of pass-through entities. The data start in 2007 and currently cover years through 2015. These data are confidential and currently can only be accessed through an internal U.S. Census project.

Key variables in Schedule K-1 include the income, deduction, and ownership share of partners and shareholders as well as the name, location, and employer identifier of the company. Unlike S-corporations, partnerships can be owned by other legal entities including partnerships and corporations. These tiered entities can make it hard to identify the ultimate owners of these enterprises when there are circular references. For more information see Goldschlag et al. (2017) and Cooper et al. (2015).

Form W-2

Form W-2 is a tax form used to report the income paid to employees in remuneration for services rendered to an employer. Employers must file a W-2 for each of their employees for services performed during the year. The availability of both employer identification numbers (EIN) and person identification numbers (SSN) allow us to identify all the salaried workers associated with employer businesses in the US. The data start in 2005 and currently cover years through 2016. Key variables in Form W-2 include the income, social security taxes, and Medicare taxes as well as individual and employer identifiers. For more information about the W2 see (https://www.irs.gov/forms-pubs/about-form-w2).

Longitudinal Employer-Household Dynamics- Employment History File (LEHD-EHF)

The LEHD-EHF is one of the core infrastructure files of the LEHD program. The EHF is sourced from quarterly unemployment insurance earnings records collected by labor market information systems across the country for unemployment insurance purposes. The EHF provides a time series of all jobs held by individuals each quarter in each state. Key variables in the EHF include employer and individual identifiers, employment quarter and year, quarterly earnings, and industry of activity. The unit of analysis is the job or an employer-employee combination. A crosswalk between the state employer identifier (SEIN) and the federal employer identifier (EIN) is available. For additional information see the LEHD infrastructure file documentation (https://lehd.ces.census.gov/doc/technical_paper/tp-2006-01.pdf).

The Annual Survey of Entrepreneurs (ASE)

The ASE is a survey of approximately 290,000 firms in the non-farm private sector. The survey is a representative sample of firms in the US with employees. Starting in 2014, the ASE was conducted on an annual basis up to 2016. The ASE will be replaced with the Annual Business Survey starting in 2017. The ASE collects information on up to 4 owners of US businesses including age, gender, race, ethnicity and veteran status. Additional information includes the business owners' education, experience, and ownership role. The ASE is the source of core demographic statistics of US business owners and includes information such as number of firms, sales and receipts, annual payroll, and employment by gender, ethnicity, race, and veteran status. The survey includes modules to collect information on specific business activities. In 2014 the ASE collected additional information on R&D and innovation and the 2015 survey asked questions about management practices. For additional information see Foster and Norman (2017).

The Census Numerical Identification System File (Numident)

The Census Numident is sourced from the Social Security Administration (SSA) applications for Social Security Numbers (Form SS-5). This is the SSA's master list of social security numbers (SSN) and includes all individuals in the US that have been issued a social security number. The Numident file is updated annually with years currently through 2016. Key variables include a protected identification key (PIK, which replaces the individual's SSN so as to protect their identity), date of birth, country of origin, gender, race and ethnicity. Starting in 1980 the SSA changed its collection of race and ethnicity so these data became non-mandatory items. The Census enhances these files with demographics data from its own data holdings including the decennial census and the American Community Survey to improve its quality.

The Patent Longitudinal Business Database Crosswalk (LPBD)

The LPBD is a crosswalk file linking individual firms to specific patents in the US Patent and Trademark Office patent grants database. Starting in 2000 the LPBD links all inventors and firms identified in patent grant documents to firms in the LBD. The LPBD uses a triangulation strategy where the best possible match is identified by comparing matches to two alternative data sources: inventors are matched to the LEHD jobs file for workers, and patent assignees are matched to the Business Register file for firms. The file starts in 2000 and is updated annually. The LPBD currently cover years through 2015. Key variables in the LPBD include firm id, patent id, application year, assignee country, assignee state, and assignee type. The LPBD is able to match in excess of 75% of all patent-assignee combinations in the USPTO and 91% of patents with US firm assignees. For additional information see Graham et al. (forthcoming).

The Private Capital Research Institute-LBD Bridge (PCRI)

The Private Capital Research Institute (PCRI) is a database of private capital data assembled by PCRI directly from several dozen private capital firms as well as from four major data vendors and private capital associations, including the Emerging Markets Private Equity Association (EMPEA), NYPPEX FUNDSIQ ("NYPPEX") Thomson Reuters, and Unquote. PCRI were matched to the Business Register using name and address linking techniques. Key variables in the PCRI database include a company id, business name, street address, zip code, state, country, day of investment, and investment category. The PCRI bridge provides a link between the LBD and the PCRI database. Match rates of US headquartered firms to the LBD are in excess of 90%. For additional information about the matching methodology see Brown and Tello-Trillo (2017). External researchers wishing to use the linked PCRI and LBD data and both internal and external Census researchers wishing to use additional PCRI variables need to submit a proposal to Leslie Jeng, Director of Research at PCRI (leslie.jeng@gmail.com). The proposal guidelines can be found at http://www.privatecapitalresearchinstitute.org/images/news/call_f_proposals.pdf.

VentureXpert

VentureXpert is a commercial database for information covering venture capital and private equity investments. The data are linked to the LBD using name and address matching techniques. Key variables include firm name and address, funding type, funding round, amounts, date of funding, and names of the VC firms. Years covered include 1980-2005.

Compustat Bridge & Compustat

The Compustat Bridge provides a link between the COMPUSTAT data and the LBD. Compustat provides financial, statistical and market information for publicly traded companies.

Prior Wage Analysis

To examine the relationship between entrepreneurial entry, success, and wages we first constructed the prior wage history of each wage earner using each individual's W-2 records. For the analysis, we defined the "prior wage" as the maximum of the annual wage payments to that individual over the prior two years (a two-year window is used to help address the timing of entrepreneurial entry, which could come mid-year).

We ran regressions to capture the entry frequency with age, both with and without controls for prior wage. The regressions take the form

$$y_{it} = age_i + \mu_t + \mu_s + \theta logwage_{it-s} + \varepsilon_{it} \quad \text{(A1)}$$

where y_{it} is equal to 1 if individual i founded a firm in year t and is 0 otherwise, age_i are age fixed effects from age 20 to 65, μ_t are founding year fixed effects, μ_s are the prior job's 4-digit industry fixed effects, and $logwage_{it-s}$ is the individual's prior period log wage. The sample consists of a randomly selected 1% of the US population from each cohort between 2007 and 2014.

Figure A5 the presents the age fixed effects, for both the regression above and for the same regression without the wage control. In explaining entrepreneurial entry, we see that the peak in middle age prevails regardless of whether we control for prior wages.

To further explore the relationship between wages and entry, and any differences for highly successful entrepreneurs, we then considered the distribution of prior wages, comparing founders with other workers. Specifically, we consider the percentile ranks of founders' wages (prior to starting their firm) in the wage distribution of the workforce.

Figure A6 shows these wage distributions. By construction, the percentile ranks for the broad workforce are uniformly distributed. Looking at all founders, we see a non-monotonicity. Founders appear disproportionately common among lower-wage workers and disproportionately common among very high-wage workers. By contrast, founders of the highest-growth firms are far more likely to come from the upper end of the wage distribution.

While descriptive, these wage results can provide further facts to discipline conceptualizations of entrepreneurship. First, individuals with quite modest outside options start lots of ordinary firms, while those with unusually strong outside options tend to start growth-oriented firms. Second, the prior wages of high-growth founders suggests these individuals have outsize success both in the labor market and in founding firms. This finding is consistent with high-growth founders being skilled in multiple domains; it is also consistent with screening, where high-growth founders set a high bar for entry into entrepreneurship, given a high opportunity cost of leaving the ordinary labor market behind.

Analysis by Calendar Cohort

While the main text pools the founding years, we can also provide additional analysis for each individual cohort year. This analysis provides a further way to generalize the findings while also demonstrating that the findings are robust outside the years 2007-2009, which overlap with the Great Recession. In particular, we provide a cohort-by-cohort analysis as far forward in time as our datasets allow. First, we extend analysis for each founding year through 2014 for our overall startup data. Second, we similarly extend the analysis for each founding year through 2014 using our "ex-ante" growth-orientation measure based on high-tech employment. Third, we extend the individual year analysis through 2011 for VC-backed startups and patenting startups, which is the limit these data allow. Finally, for "ex-post" growth outcome measures, and shortening the post-founding window to three years, we can look at individual cohorts through 2011. Table A4 presents the average founder ages for these separate cohorts. We see that the middle-age tendency is highly robust.

Appendix II: Additional Figures and Tables

Figure A1: Founder Rates by Age

Fig. A1-A: Size of Workforce by Age

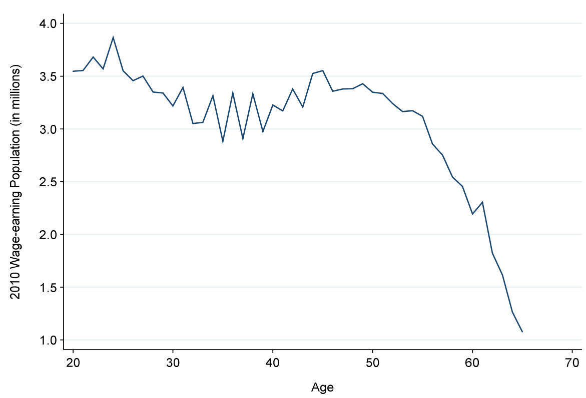


Fig. A1-B: Founders per Worker, by Age

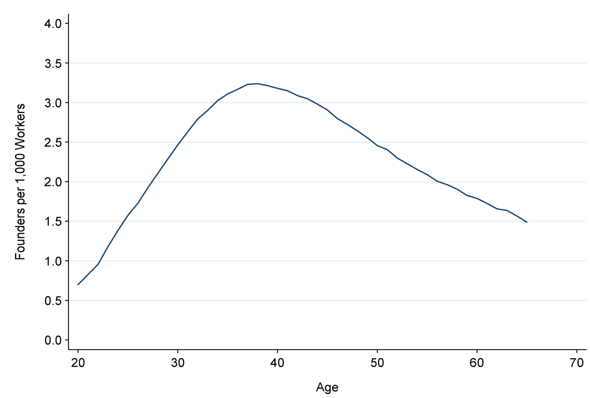
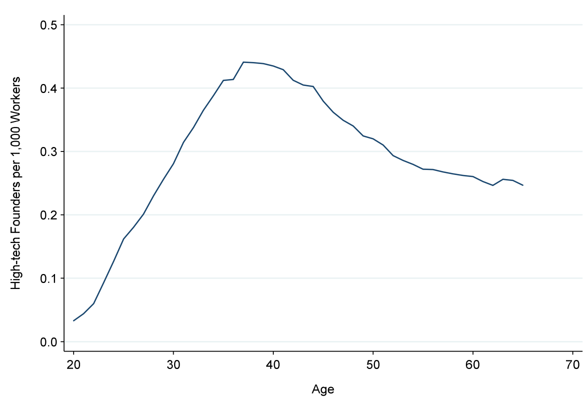


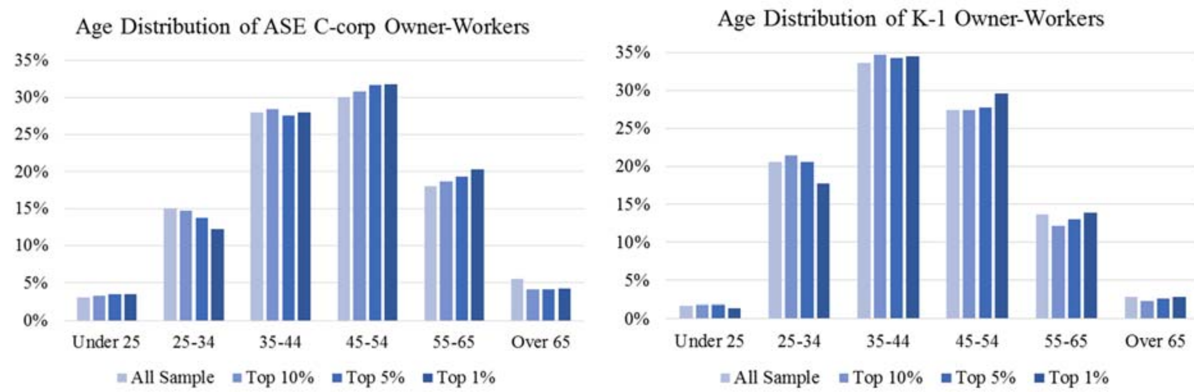
Fig. A1-C: Tech Founders per Worker, by Age



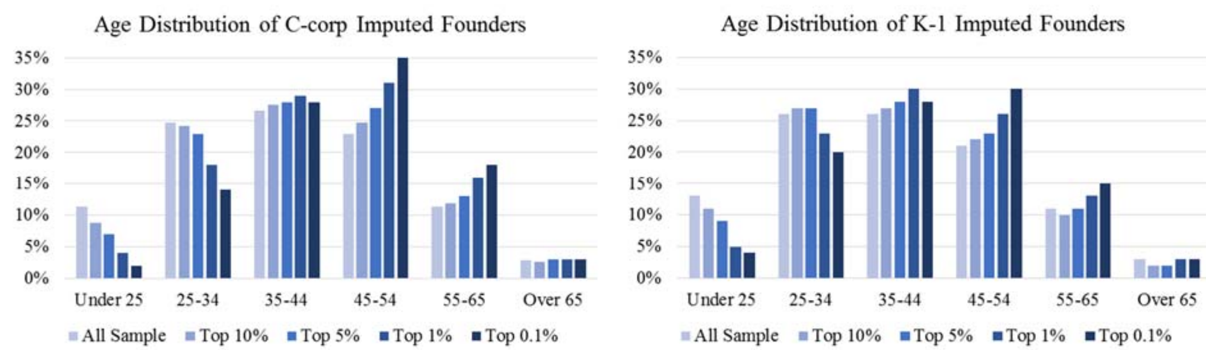
Source: Authors calculations based on W-2 earnings records, form K-1, and Longitudinal Business Database between 2007 and 2014.
Notes: These figures show the number of wage earners, founders, and high-tech founders in the US. Each bin represents an age cohort. Ages between 20 and 65 are incorporated in the plots. Figure A1-A uses the 2010 W-2 file. Figures A1-B and A1-C use data over 2007-2014.

Figure A2: Results by Founder Definition and Legal Form

Panel A: Owner-Worker, C-Corporation and K-1 Firms



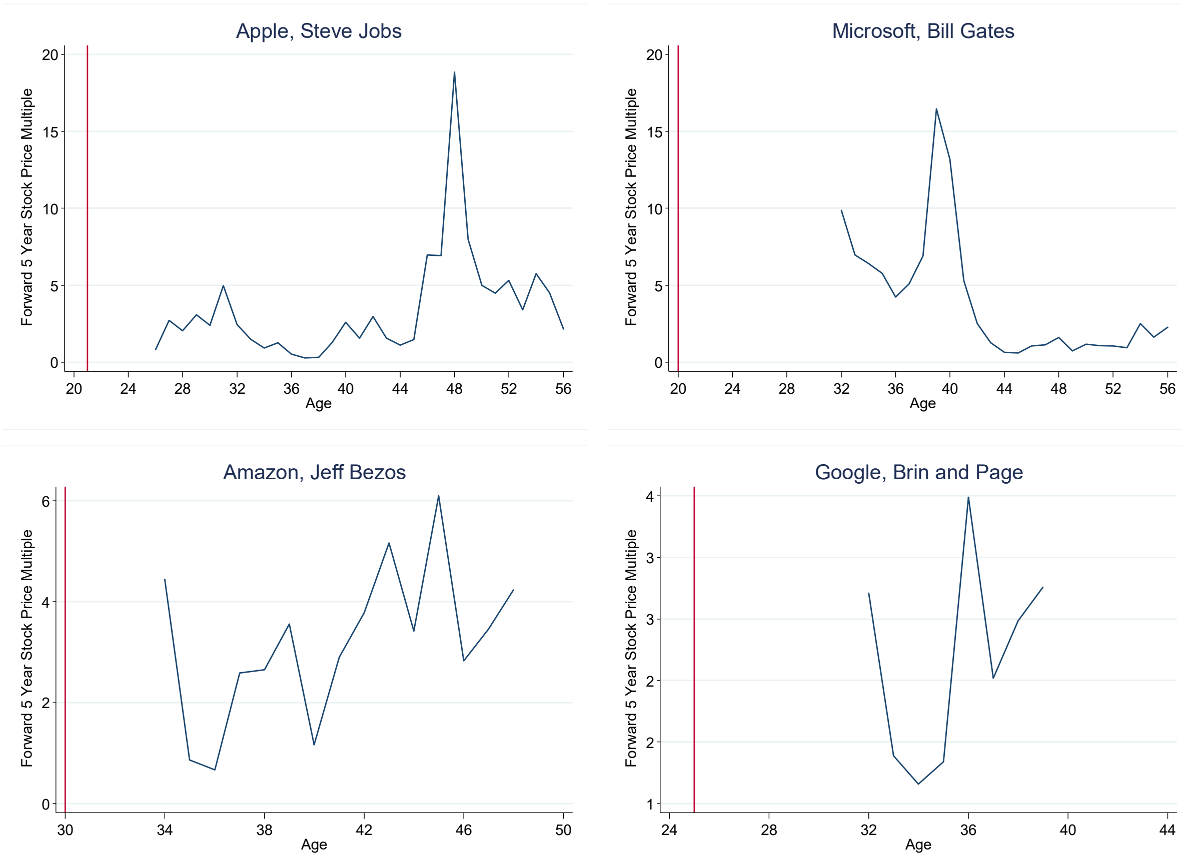
Panel B: Initial Top 3 Earners, C-Corporation and K-1 Firms



Source: Authors calculations based on Longitudinal Business Database, W-2 earnings records, form K-1, and Annual Survey of Entrepreneurs.

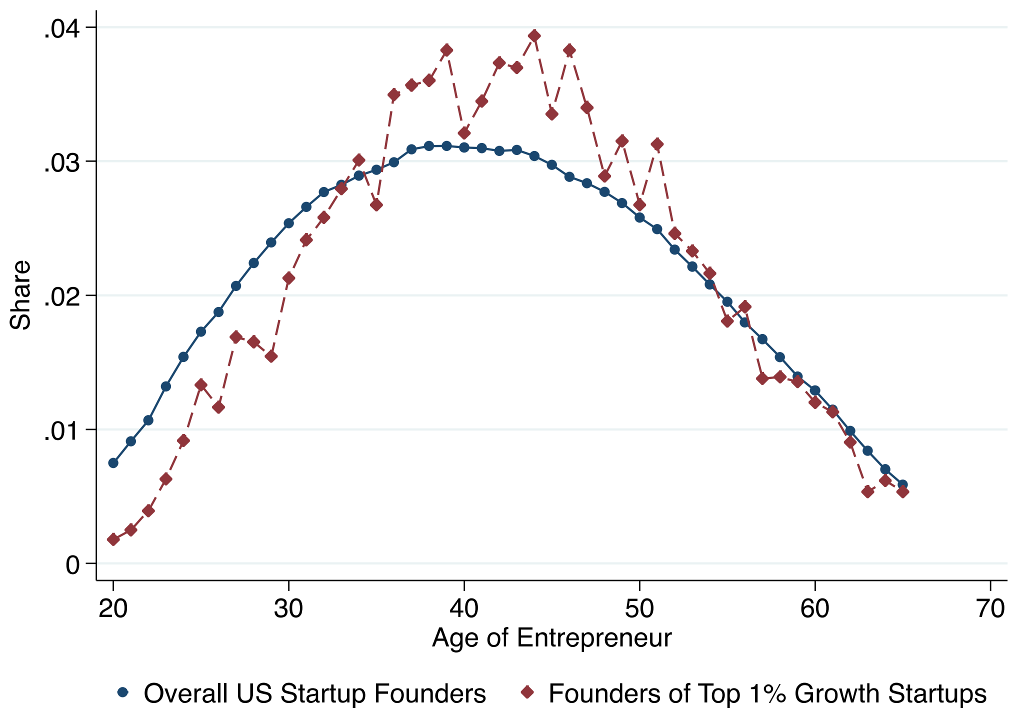
Notes: Startup firms born between 2007 and 2012 in the Annual Survey of Entrepreneurs (ASE) are included for the left side of Panel A. Growth outcomes are calculated over a three-year window for each cohort and the top 1%, 5% and 10% is identified from the distribution. The rest of the figures include all new C-corporations, S-corporations, and Partnerships in the Longitudinal Business Database (LBD) born between 2007 and 2011. Growth outcomes are calculated over a three-year window for each cohort and the top 0.1%, 1%, 5% and 10% is identified from the distribution. The left side of Panel A is based on founders of C-corporation firms in the Annual Survey of Entrepreneurs. The right side of Panel A is based on founders of S-corporations and partnerships in the K-1 database. Panel B is based on imputed founders (first-year joiners who are among the top three earners) using W-2 wage-records.

Figure A3: Forward Stock Multiples as the Founder Ages:
Apple, Microsoft, Amazon, and Google



Source: Authors calculations based on public data.
Notes: The vertical red line indicates the founders' age in the year of the firm's founding, and the x-axis presents the age of the indicated founder as time passes. The forward stock-price series begins in the year of the initial public offering for each firm. For Google, Brin and Page were born in the same year (1973). Historical share prices are sourced from Bloomberg.

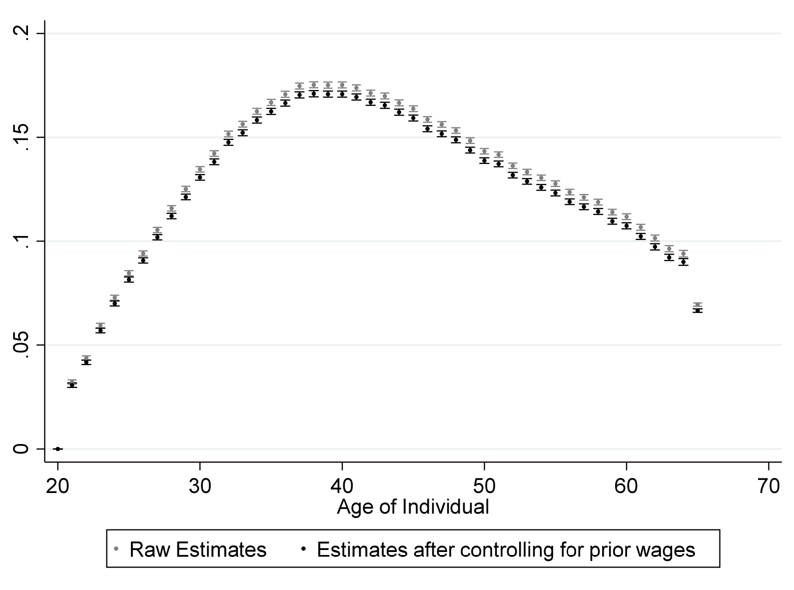
Figure A4: Founder Age Distribution:
All Startups and High Performance Startups by Sales (5 Years after Founding)



Source: Authors calculations based on W-2 earnings records, form K-1 and Revenue-Enhanced Longitudinal Business Database.

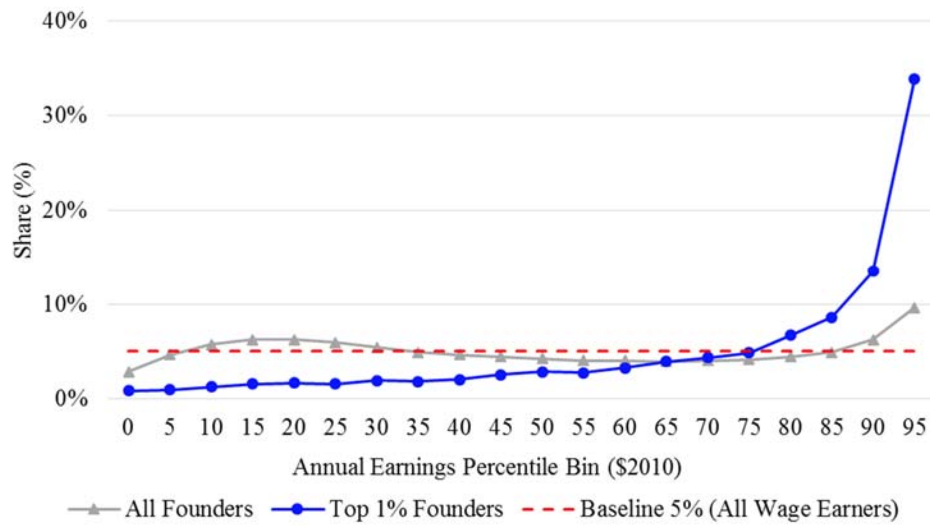
Notes: This set of kernel density plots shows the age distribution of startup founders (at year of founding) in the US. Each bin represents an age cohort. Ages between 20 and 65 are incorporated in the plots. The blue (left) plot incorporates all founders of new C-corporations, S-corporations, and Partnerships with employees founded between 2007 and 2014 as identified in the Longitudinal Business Database (LBD). The red (right) plot represents founders of the top 1% growth firms founded over the 2007-2008 period, given that revenues data are available up to 2013. Top 1% revenue growth threshold value is calculated for each yearly cohort based on the real revenue figures from the LBD in the five years after the birth of the firm.

**Figure A5: Founder Age Distribution,
With and Without Controls for the Founders' Prior Wages**



Notes: This figure presents estimates of the age indicator variables in the regression equation (A1), together with their associated 95% confidence intervals, with and without prior wage controls for the individuals in the sample, which consists of a randomly selected 1% of the US population of wage earners in the W-2 from each cohort between 2007 and 2014.. See online appendix text for details of the data construction and regression specification.

Figure A6: Wage Distributions of Non-Founders, Founders, and Successful Founders



Notes: This figure examines the wage distribution of founders and highly successful founders compared to the background wage distribution of the workforce. Prior wages are calculated from W-2 records and translated into 2010 U.S. dollars. The x-axis represents percentile bins of annual earnings. By construction, the percentile rank for the workforce as a whole is uniformly distributed. Top 1% founders are those whose firms achieve top 1% employment growth within 3 years.

Table A1: Founder Age – Perceptions from Media & Two Prominent VCs

	TechCrunch Awards	Inc. and Entrepreneur Magazines	Sequoia	Matrix Partners
Mean	31.0	29.1	33.9	36.5
Median	30	27	33	36
(St. Dev.)	(7.1)	(7.0)	(8.7)	(8.6)
Observations	232	51	415	246
Period	2008-2016	2015	1969-2014	1948-2014
Sectoral Focus (top 5)	Education, Software, Social Media, Consumer Electronics, e-Commerce	Technology, Retail, Media, Consumer Goods, Food Delivery	Semiconductors, Networks, Task Mgmt. Apps, Website Compilers, Cloud	Networks, Applications, Commerce, Platform/ Infrastructure, Semiconductors/ Materials

Notes: TechCrunch gives annual awards to the "most compelling startups, internet and technology innovations of the year". Inc. magazine and Entrepreneur magazine provided "Entrepreneurs to Watch" lists in 2015. The founder ages for new ventures backs by the two venture capital firms (Sequoia Capital and Matrix Partners) were obtained by the authors through researching all the companies listed on their respective websites.

Table A2: Summary of Data Sets

Dataset	Units and coverage	Relevant Variables	Period and Frequency	Access
Longitudinal Business Database (LBD)	• Establishments and firms • All private non-farm employers in the US and outlying territories	Firm identifier, establishment identifier, payroll, employment, industry, location, legal form of organization	Annual, 1976-2015	FSRDC/Census approved projects
Schedule K-1 (Form 1065/1120)	• Owners • All pass through entities (partnerships and S-corporations)	Individual identifier, firm identifier, business income, deductions, share of profit/loss	Annual, 2007-2016	Census Bureau employees on approved projects and a need to know
Form W-2	• Employees • All workers in the US for whom employers made payments	Individual identifier, employer identifier, wage income, social security, or Medicare wages.	Annual, 2005-2016	Census Bureau employees on approved projects and a need to know
Longitudinal Employer-Household Dynamics-Employment History File (LEHD-EHF)	• Salaried workers by employer • All salaried workers subject to unemployment insurance	Individual identifier, employer identifier, earnings (quarterly and annualized), industry	Quarterly, 20XX-2015 (Initial year varies by state)	FSRDC/Census approved projects
Annual Survey of Entrepreneurs (ASE)	• Businesses • Sample of 290,000 non-farm businesses with paid employees	Firm identifier, information for up to 4 owners including age, gender, race,	Annual, starting in 2014-2016 to be replaced by	FSRDC/Census approved projects

	and receipts of $1,000 or more	ethnicity, education, experience and type of ownership	the Annual Business Survey in 2017	
Census Numident	• Individuals • All individuals with a US social security number	Individual identifier, date of birth, gender, race, ethnicity, country of origin	Updated annually	FSRDC/Census approved projects
Longitudinal Patent Business Database (LPDB)	• Patent-firm links • All patents in the USPTO grants database matched to the LBD	Firm identifier, Patent identifier, year	Annual, 2000-2014	FSRDC/Census approved projects
Private Capital Research Institute-LBD Bridge (PCRI)	• Firms • Private capital deals including buy outs, VC, growth equity, secondary purchases.	Firm identifier, Category of private capital	1990-2015	FSRDC/Census approved projects prior approval of PCRI
VentureXpert	• Firms • VC deals	Firm identifier, Venture capital funding	1980-2005	Data provided by researcher through a license agreement
Compustat-Bridge	• Publicly traded firms	Firm identifier, financial and market data	1976-2013	FSRDC/Census approved projects prior approval of PCRI

Table A3: Founder Age — Oldest and Youngest Technology Sectors

Panel A: Technology Sectors, Youngest 5

NAICS Code	Sector	N	Mean
5172	Wireless Telecommunications Carriers (except Satellite)	1,500	38.5
5182	Data Processing, Hosting, and Related Services	6,100	39.7
5112	Software Publishers	3,600	39.8
5415	Computer Systems Design and Related Services	100,000	40.1
8112	Electronic and Precision Equipment Repair and Maintenance	4,900	40.8

Panel B: Technology Sectors, Oldest 5

NAICS Code	Sector	N	Mean
4862	Pipeline Transportation of Natural Gas	50	51.4
3251	Basic Chemical Manufacturing	700	47.9
3255	Paint, Coating, and Adhesive Manufacturing	400	47.5
2111	Oil and Gas Extraction	3,100	47.5
3336	Engine, Turbine, and Power Transmission Equipment Manufacturing	400	47.3

Notes: Sector is shown in the first column, observation counts of founders in the second column, and mean founder age in the third column. Sectors are defined at the 4-digit NAICS level. Only new firms are included. Counts are rounded to comply with disclosure rules. Sample is all new businesses in the U.S. from 2007-2014 based in the Longitudinal Business Database (LBD).

Table A4: Mean Founder Age by Calendar Year of Firm's Founding

		Ex-Ante Startup Type			Ex-Post Startup Success	
Founding Year	All Startups	High-Tech Sectors	VC-backed Firms	Patenting Firms	Top 1% (3-yr)	Successful Exits
2007	41.8	43.2	42.4	44.0	43.8	46.3
2008	41.8	43.2	42.2	44.2	44.2	46.2
2009	41.8	43.3	42.7	45.2	44.6	46.1
2010	41.8	43.4	41.6	45.0	44.1	46.9
2011	41.8	43.4	41.5	45.3	44.9	47.5
2012	41.8	43.1	-	-	-	-
2013	42.0	43.0	-	-	-	-
2014	42.5	43.3	-	-	-	-

Notes: This table presents the mean age of founders by year of founding (rows). Mean age is presented subject to data availability of the growth-orientation measure (columns). Data for all new ventures and for new ventures in high-tech sectors are available through 2014. VC-backing and patenting firms are known for firms in the LBD through 2011. For ex-post growth performance, employment growth uses a 3-year window to determine upper tail firms. This growth measure and the successful exit measure are known for new ventures starting through 2011.

Table A5: Minimum and Maximum Ages within Founder Teams

Panel A: Owner-Worker Definition of Founders (K-1)

	All Startups	High-Tech Startups	VC-backed Startups	Patenting Startups	Top 1%	Top 0.1%	Successful Exit
Within Startup							
Min Founder Age	42.7	44.0	39.8	43.6	40.9	42.3	43.3
Max Founder Age	44.6	45.5	47.8	46.9	45.6	47.8	47.1

Panel B: Initial Team Definition (K-1 and C-Corporations)

	All Startups	High-Tech Startups	VC-backed Startups	Patenting Startups	Top 1%	Top 0.1%	Successful Exit
Within Startup							
Min Founder Age	35.1	39.1	36.5	37.8	35.0	37.4	38.5
Max Founder Age	46.0	45.7	47.3	48.4	50.1	51.4	51.4

Notes: Panel A incorporates all S-corporations and Partnerships founded over the 2007-2014 period in the Longitudinal Business Database (LBD), except for the Top 1% and Top 0.1% columns, which include those firms founded over the 2007-2009 period for which we can observe 5 years of employment data after founding. Panel B incorporates all S-corporations, Partnerships, and C-corporations founded over the 2007-2014 period, except for the Top 1% and Top 0.1% columns, which include those firms founded over the 2007-2009 period for which we can observe 5 years of performance data after founding.

References

Brown, David and Cristina Tello-Trillo. “The PCRI/LBD Bridge File.” Mimeo, 2017.

Cooper, Michael, John McClelland, James Pearce, Richard Prisinzano, Joseph Sullivan, Danny Yagan, Owen Zidar, Eric Zwick (2015), “Business In The United States: Who Owns It And How Much Tax Do They Pay?” Tax Policy and the Economy 2016 30:1, 91-128

Goldschlag, Nathan, J. Daniel Kim, and Kristin McCue (2017), "Just Passing Through: Characterizing U.S. Pass-Through Business Owners." Working Papers 17-69, U.S. Census Bureau Center for Economic Studies Working Paper No. 17-69.

Graham, Stuart J. H., Cheryl Grim, Tariqul Islam, Alan C. Marco and Javier Miranda (forthcoming), “Business Dynamics of Innovating Firms: Linking U.S. Patents with Administrative Data on Workers and Firms.” Journal of Economics and Management Strategy.

Jarmin, Ron and Javier Miranda (2012), “The Longitudinal Business Database”. Center for Economic Studies, CES-WP-02-17.