Founder-CEO Compensation and Selection into Venture Capital-Backed Entrepreneurship*

Michael Ewens
Columbia Business School

Ramana Nanda
Imperial College London

Christopher Stanton
Harvard Business School

Forthcoming at the Journal of Finance
September 2023

Abstract

We show theoretically that a critical determinant of the attractiveness of VC-backed entrepreneurship for high-earning potential founders is the expected time to develop a startup's initial product. This is because founder-CEOs' cash compensation increases substantially after product development, alleviating the non-diversifiable risk that founders face at startup birth. Consistent with the model's predictions of where the supply of entrepreneurial talent is likely to be most constrained, we find that technological shocks differentially altering the expected time to product across industries can explain changes in both the rate of entry and characteristics of individuals selecting into VC-backed entrepreneurship.

Keywords: Entrepreneurship, venture capital, executive compensation.

JEL Classification: G24, G32, J33.

*We are especially thankful to Bob Hall and Susan Woodward for sharing their code. We are also grateful to Ulf Axelson, Zoe Cullen, Matthew Denes (discussant), Roberta Dessi, Daniel Ferreira, Paul Gompers, Juanita González-Uribe (discussant), Stephen Hansen (discussant), Victoria Ivashina, Steve Kaplan (discussant), Francisco Queiro (discussant), Ben Roth, Kathryn Shaw (discussant), Pian Shu (discussant), Peter Thompson, Ting Xu and participants at presentations at the AFA, Caltech, Carnegie Mellon, CEPR, CUHK, CityU Hong Kong, FIRS, Georgia Tech, HBS, HEC Paris Entrepreneurship Workshop, Imperial, London School of Economics, Nanyang Technological University, NBER Summer Institute, NOVA SBE, Michigan, University of Virginia, UT Dallas and the WFA for helpful discussions. Nanda and Stanton thank HBS for financial support. Ewens has advised and received financial support from Correlation Ventures. Correlation Ventures provided some of the data, but did not have the right to review results. Advanced HR, which provided data on founder-CEO compensation, has reviewed the paper for disclosure of individually identifiable information but did not have the right to review content.

1 Introduction

Venture capital (VC)-backed startups play a critical role in the economy due to their disproportionate contribution to corporate R&D and innovation (Kortum and Lerner 2000; Lerner and Nanda 2020). While a growing literature studies frictions in the provision of VC financing to startups seeking risk capital, very few papers investigate potential constraints to the *supply of entrepreneurial talent* to VC-backed startups — and in particular to the drivers of selection into VC-backed entrepreneurship.

Understanding selection into VC-backed entrepreneurship is critical to understanding startup innovation because, as noted by Rajan (2012), the differentiation needed to create a new enterprise also requires an entrepreneur to hold significant control rights over its direction. This requirement implies that an entrepreneur cannot be decoupled from the startup at its birth. Given the systematic association between individual backgrounds or experiences and entrepreneurial ideas (Gompers, Lerner, and Scharfstein 2005; Koning, Samila, and Ferguson 2021), *who* finds it feasible to select into entrepreneurship has clear implications for the types of innovations commercialized by startup ventures.

A potential barrier to the supply of entrepreneurial talent stems from the compensation contract between investors and entrepreneurs at firm birth (Hall and Woodward 2010; Rajan 2012). Since this contract comprises low cash compensation to screen entrepreneurs, it can also leave entrepreneurs, especially those with attractive outside options, holding substantial non-diversifiable risk at founding. Yet, as discussed by Rajan (2012), startups undergo a transformation during their lifecycle such that soft assets like founder human capital decline in importance, enabling the "standardization" that facilitates financing from arms-length investors. If founder-CEOs' compensation changes after this transformation, understanding when this transformation occurs in a startup's lifecycle is critical, as it can have important implications for the degree of non-diversifiable risk held by entrepreneurs. In turn, this can impact the degree to which the supply of entrepreneurial talent is constrained.

This paper combines several novel data sources with a dynamic model of selection into entrepreneurship to study the length of time founder-CEOs hold non-diversifiable risk and the implications this has for the supply of entrepreneurial talent to VC-backed startups. We pinpoint the development of a startup's first product as the key milestone after which CEOs receive substantial increases in cash compensation, consistent with a shift in liquid pay over the "differentiation" to "standardization" transformation modeled by Rajan (2012). Indeed, we show theoretically that increases in CEOs' cash compensation after a startup begins product development are large enough that non-diversifiable risk held by founder-CEOs falls substantially after achieving the product development milestone. In turn, this implies that a startup's expected time to develop its first product is a key determinant of the attractiveness of VC-backed entrepreneurship. Our empirical analysis provides strong support for the theoretical predictions. In doing so, we offer some of the first evidence for the particular sectors where the supply of entrepreneurial talent is most likely to be constrained, and hence where nascent innovations are most likely to be missing.

Our analysis proceeds in three steps. We first use proprietary data on CEO compensation for a representative sample of U.S. VC-backed startups to study how founder-CEO compensation evolves across a startup's lifecycle. As might be expected with a typical "entrepreneur compensation contract," cash compensation for founder-CEOs at the very early stages is comprised of a low salary with no bonus. However, we find that developing a first product is an important milestone associated with a clear shift in CEO cash compensation. Baseline salary increases markedly after this milestone, and the structure of cash compensation is more consistent with typical "manager compensation contracts," as documented in the literature on CEO compensation in publicly-traded firms (Edmans, Gabaix, and Jenter 2017). The product development milestone is also the point at which the hazard of founder-CEO replacement substantially increases. After product development, the firm and the entrepreneur can be decoupled from each other.

An important aspect of our findings is that this shift to a "manager cash compensation contract" is also present for founder-CEOs who remain with the venture. To the extent that the increased cash compensation post-product can insure an entrepreneur against non-diversifiable risk, the expected time to product development can have important implications for the attractiveness of selecting into entrepreneurship. In the second step of our analysis, we quantitatively assess the attractiveness of entrepreneurial entry as a function of the expected time to product development. To do so, we use a dynamic consumption-savings model in which founder-CEO cash compensation responds to product development milestones, as estimated from the data on startup CEO compensation. While we find that post-product increases in cash compensation make founding a startup more attractive than a contract that pays a flat salary, low pre-product compensation can still leave high-earning potential founders holding substantial non-diversifiable risk. This result implies that variation in the expected time to product development is a key driver of the relative attractiveness of entrepreneurship across sectors for potential founders with attractive options outside of entrepreneurship and insufficient personal wealth to finance consumption while waiting for a product to materialize.[1]

Our model predicts that industries where the product milestone is expected to be achieved more quickly, and hence uncertainty is resolved faster, are those industries where entrepreneurship is *ex ante* significantly more attractive. Conversely, all else equal, a missing supply of entrepreneurs is more likely in industries with high expected time to product development, leading uncertainty to be resolved slowly and entrepreneurs left bearing the burden of non-diversifiable risk.

We illustrate that the model-implied differences in founding attractiveness are quantitatively important by using the introduction of cloud computing as a shock to product development times. This 2006 shock disproportionately affected software and internet startups. A parametric model of the joint distribution of income and wealth shows that founding becomes attractive for many more high-earning individuals. Prior to the advent of cloud computing, our model suggests that non-diversifiable risk potentially deterred 14% of individuals in the top 2% of the earnings distribution in the United States (i.e., with pre-tax earnings above $225,000) from founding a VC-backed startup. The introduction of cloud computing – which reduced the share of three year-old firms that had yet to transition to having a product by 15% – made founding attractive for an additional 61,000 individuals with earnings over $225,000. This increase represents about 13 percent of the pre-cloud era individuals deterred from entering by non-diversifiable risk and about 2 percent of the total population earning more than $225,000. To put these numbers into context, there were about 1,100 VC-funded startups founded annually across all industries between the early years of cloud computing (2001 to 2006). Of course, not all of these potential founders will have an idea or the motivation to start a company. Nonetheless, our analysis suggests that differential changes in expected time to product across industries impacts the composition of those sectors' high-earning entrants.

[1]While our model builds on Hall and Woodward (2010), they assume a fixed, low salary through the life of the startup. Their model does not consider variation in the compensation contract that arises with product milestones. We explicitly model the estimated time to product as a key parameter that can impact the decision to select into entrepreneurship, which we find is quantitatively important.

In the final part of the analysis, we test the model's predictions using LinkedIn data on the career histories of entrepreneurs whose startups received venture capital financing. A novel aspect of this analysis is our use of pre-founding job titles and industries to match to occupation and industry codes in the Census Bureau's American Community Survey (ACS) - which has information on earnings for a representative sample of the US Population. This allows us to examine how the pre-entry earnings of VC-backed entrepreneurs vary based on whether they founded their venture in industries with longer or shorter expected time-to-product.

To overcome concerns about systematic differences in skill requirements or founders' motivations across sectors, we exploit the cloud computing technology shock in difference-in-differences regressions. The number of founders that we estimate to have earned more than $225,000 prior to founding increases by over 60% in internet, software, and related entries after the advent of cloud computing. This greater entry into VC-backed entrepreneurship from the top of the earnings distribution was only seen in industries where the time to product fell after the diffusion of cloud computing. There was little change in the composition (or rate) of entrants in industries where the time to product did not change. These results provide strong empirical evidence consistent with shorter expected time to product development playing a critical role in relaxing entrepreneurial entry constraints for the very highest earning individuals in the workforce, with implications for the supply of new ideas in these industries.

Our findings relate to several strands of literature. First, our paper is relevant to the literature on principal-agent problems, both between investors and entrepreneurs and between shareholders and CEOs. On the one hand, this work has highlighted the importance of contracts that provide minimal cash compensation to help screen entrepreneurs under asymmetric information (Aghion and Bolton 1992). On the other, it has highlighted the importance of dynamic contracts where performance is rewarded with intermediate cash liquidity (Edmans et al. 2012). Our work shows that the transition from one type of contract to another occurs relatively early in the life of venture capital-backed startups, and consistent with Rajan's (2012) theoretical model of the lifecycle of firms, is tied to the process of standardization when the founder and the firm can effectively be decoupled from each other.

Our paper is also related to the literature on selection into entrepreneurship. Past work has documented the role of paid employment as a source of ideas and training for potential entrepreneurs and the conditions leading them to select into entrepreneurship (Bhidé

2003; Gompers, Lerner, and Scharfstein 2005; Babina, Ouimet, and Zarutskie 2017; Kim 2018). Theoretical work has examined the financial trade-offs between entrepreneurship and paid employment (Anton and Yao 1995; Hellmann 2007), as well as the potential frictions associated with the entry decision in terms of non-diversifiable risk (Hall and Woodward 2010). We show that for VC-backed startups, expected time to product development (rather than expected time to exit) is likely the core driver of the degree of non-diversifiable risk faced by entrepreneurs. This fact is consistent with conceptualizing entrepreneurial selection as an experimentation process, where the speed with which milestones are achieved, and hence uncertainty resolved, plays an important role in determining the extent of risk facing entrepreneurs (Kerr, Nanda, and Rhodes-Kropf 2014; Manso 2016; Dillon and Stanton 2018).

Finally, our results highlight an under-appreciated role played by venture capital investors– as intermediate liquidity providers – which they might be uniquely positioned to do as hands-on investors who are able to resolve information asymmetry more effectively than passive capital providers. This liquidity provider role may also help explain the sectors where VCs are more actively involved in financing innovation. The faster resolution of uncertainty for firms in certain sectors (Ewens, Nanda, and Rhodes-Kropf 2018) implies a lower need for VCs to provide up-front inducements to founders. Sectors where uncertainty is resolved more slowly likely exhibit the highest burden of non-diversifiable risk. The degree to which these individuals' ideas are not commercialized (or commercialized inside incumbent firms) as well as the aggregate impact of this selection remains an interesting area of further work.

2 Startup's First Product as a Key Inflection Point

This section documents the results from the first step of our three-step analysis. We show that transitioning to product development is an important milestone for a startup, marking the point where founders can be decoupled from their entrepreneurial ventures. The hazard of founder replacement increases substantially after a firm makes this transition. Moreover, all CEOs, including founder-CEOs who remain with the firm, get substantial increases in liquid cash compensation after this milestone, suggesting a shift from an 'entrepreneur contract' to a 'manager contract' for the founder-CEO once a firm has begun product development. We first document results for CEO compensation and then for founder-CEO replacement.

2.1 Founder-CEO Compensation

2.1.1 Description of Compensation Data

Our core compensation dataset comes from two cross sections of compensation surveys from Advanced HR (AHR), a leading data provider covering private startups. Startups become eligible for survey inclusion by AHR if they have received investment from a venture capital investor that participates in the survey. Participating investors range from very early-stage seed funds to later stage funds. Nearly all of the most prominent and well-known VC funds are included. Based on data from VentureSource, a commercial data provider that contains information on the near-population of VC-backed startups, participating funds in the AHR survey managed 42% of total industry assets and deployed nearly 49% of the dollar-weighted investments during the survey years.[2]

Survey completion happens after a participating VC fund solicits portfolio company responses. Responding to the survey by startup portfolio companies is strongly encouraged by the investors, who are typically members of the startup's board. Both VCs and startups get access to benchmarking data in exchange for responses. We are aware of no other compensation data for startups that offers similar coverage in our sample period or on the scale that AHR provides.

Each survey wave contains individual-level information on salary, bonus, fully diluted equity, and an indicator for whether the CEO is a founder. The individual-level records also include a number of coarsened, categorical startup-level characteristics, such as revenue, total employment, cumulative venture financing raised, and product-related milestones. To protect confidentiality, our data and that shared with venture capital investors are anonymous. The data cannot be linked to individuals or startups, and startups cannot be linked over time. However, because we know the identities of the VC firms that participate in the AHR survey, we use the VentureSource data to estimate the implied response rate among their portfolio companies. We also detail potential differences between respondents and startups that were eligible to take the survey. We discuss these analyses in the next section on Survey Data Validation.

We have access to data from AHR's 2015 and 2017 survey waves for technology companies

[2]Appendix A.1 provides details about how we compare the samples and Appendix B.1.1 details the creation of the VentureSource data for this benchmarking exercise.

(which does not include data on biotech and healthcare startups). The 2015 survey contains data from 933 portfolio companies that received funding from 70 VC investors; the 2017 survey has data on 1,552 portfolio companies associated with 115 venture capital investors.[3] Our core AHR sample focuses on U.S.-based CEOs in startups founded after 1996.[4]

We also use three additional sources to verify the robustness of our compensation: data from Compensation Pro, that conducted a survey of VC-backed startup in 2008–2010; S-1 filings of startups that go public to study the cash compensation of founder-CEOs of newly public firms; and data on compensation for CEOs of publicly traded firms using Execucomp. None of these data sources has representative coverage of early stage firms, so we cannot use them for inference on compensation dynamics over the firm lifecycle. Instead, these data allow us to show that the compensation patterns that we document in the AHR data for mature, post-product firms are similar to those from prior periods and across different industries, including startups in the biotech, consumer products and industrial/energy sectors. This result provide reassurance that the relatively high post-product cash compensation we see in the more detailed AHR data is not being driven by the particular time period of our analysis.

2.1.2 Validating AHR's Compensation Data

Although the survey data are anonymous, we know the identities of the participating VCs, which enables us to validate the AHR surveys with comparisons to the universe of startups active during the survey years, as measured by VentureSource. By identifying startups in the VentureSource universe with at least one VC investor that participates in the survey, we calculate that the AHR data covers 58% of the VentureSource firms backed by eligible VCs, indicating a strong response rate among startups eligible for survey inclusion.

The AHR data covers 25% of all VentureSource firms that are under 10 years old in the industries it covers, which is consistent with an approximately 60% response rate among

[3]The increase in the number of portfolio companies arises largely from the increase in investors who participate in the 2017 wave, including earlier-stage seed funds, and corporate venture arms.

[4]We drop 24 startups that are listed as having only growth capital or that have received over six rounds of funding. We exclude 38 clean tech companies from the 2015 wave and 35 from 2017, as clean tech investment was relatively small for VCs during our sample nor was prominent before or after. Our results are insensitive to including these firms and are instead conservative, as clean tech firms have the largest change in median pay between pre- and post-product status. We also exclude startups if we do not know the location of the CEO, eliminating two observations.

eligible startups, and the VCs who are part of AHR accounting for about 40-45% of all venture capital investments in the U.S. Among firms with investments from participating VCs, the AHR sample appears representative based on total investment amounts and the firm age distribution. We do find that the AHR data appears to have a somewhat higher response rate from firms that have raised more than the median level of capital, so in later analyses we re-weight the AHR survey data to reflect the joint distribution of firm age and capital raised in the VentureSource universe; these results are very similar to un-weighted specifications. Appendix A.1 provides details about how we compare the samples, the calculations we conduct, and tables displaying additional moments for comparison.

2.1.3 Descriptive Statistics for CEO Compensation in VC-backed Startups

Table 1 displays descriptive statistics for founder-CEOs in the AHR data. The data are presented in two panels based on variables that capture financing information and performance milestones. The first panel splits by funding round, capturing the number of rounds of outside investment the firm has raised. The next panel presents data broken down for firms based on revenue and product status. Both panels are sorted from earlier to later stages for each variable, which correlates with increasing average firm age across rows. Note that this progression is not deterministic because there is variation in firm age in each category. This variation allows separation of age effects from other milestones or events that influence compensation over the startup lifecycle.

Average CEO cash compensation increases with the company stage, with the largest percent increase in compensation occurring between "Pre-Product" and "Post-Product" firms. Pre-Product firms are those that report "Early / Product Definition" as their development stage. Post-Product firms report "Product Development, "Product in Beta", "Shipping Product" or "Profitable" as their product development stage. All firms with strictly positive revenue are "Post-Product" firms although it is possible to be categorized as Post-Product but Pre-Revenue. Panel B shows a dramatic difference in compensation between post-product firms that have yet to realize any revenue and pre-product firms. Founder-CEOs of pre-product firms earn $85,000 per year in cash, whereas CEOs of post-product, pre-revenue firms average $165,000 in annual cash compensation.

Progressing through rows in each panel of Table 1, CEO compensation rises with each stage or milestone. The CEOs of mature firms, (Post-Series B, with greater than $10 million in revenue), earn, on average, over $240,000 per year in cash compensation. The standard deviation of total cash compensation also increases across rows in each panel.[5] The last column of the table shows that CEO's hold less than 40% of the as-if-converted equity by the first financing round.[6] Founder equity stakes decline as firms mature due to the dilution that comes with raising capital. Our analysis of the attractiveness of founding will account for the increasing profile of pay and firm value with respect to milestones and founders' declining equity shares that comes from dilution.

Graphical evidence clearly displays the importance of having a product as a pre-condition for increases in founder-CEO compensation. The top left panel of Figure 1 displays average cash compensation and the interquartile range by firm age.[7] The figure shows a substantial increase in founder-CEO cash compensation as firms age. The remaining panels in Figure 1 help to disentangle mechanisms. The top right panel conditions on "Pre-Product" firms that have no revenue and have not yet achieved viable product definition. Across the firm age distribution, average cash compensation is below $100,000 for pre-product firms. Founder-CEOs of pre-product firms do not have an increasing compensation profile with firm age. Founder-CEOs of four year old pre-product firms earn no more than one year old pre-product firms. Note that the distribution of pre-product firms is not truncated at four years of age—instead, firms that do not have a product rarely survive to their fifth year. This pattern suggests that the overall increase in compensation with firm age stems from the increasing number of surviving firms that have achieved product market milestones and the death of firms that fail to achieve product traction.

The bottom two panels plot similar figures, but instead of focusing on firm age, the x-axis is capital raised. Capital is related to firm milestones, yet conditional on capital raised, pre-product firms in the bottom right panel have lower average pay and a more compressed interquartile range of pay than the unconditional plots. Both of these splits, by firm age and

[5]Note that while the coefficient of variation is high for pre-product firms, this is largely driven by a subset of CEOS of very early-stage, pre-product companies who take de-minimis salaries. For about 15% of these firms, the CEO earns less than $20,000 per year. Thus, the variation at very early stages comes from CEOs who earn significantly less than the mean.

[6]The CEO is likely one of two to three founding team members that will collectively own more than the reported number in Table 1. See Hellmann et al. (2019) for an analysis of within-founding team ownership dynamics.

[7]Note that this analysis is conditional on surviving firms, but surviving firms are the relevant sample for assessing founder risk. Upon firm failure or an executive's exit, he or she earns their outside compensation. We later assess whether startup experience itself changes the outside option relative to other career paths.

capital raised, suggest that having a viable product is a significant inflection point for CEO compensation. We now turn to multivariate regressions to demonstrate that these insights remain when we account for other factors that may move with product status. Moreover, a bounding exercise along the lines of Lee (2009) that accounts for potential non-random survey participation by product status and firm age shows that our results are not due to unobserved selection.

2.1.4 Regression Evidence

Table 2 displays Poisson regressions where the dependent variable is total cash compensation for the founder-CEOs of VC-backed firms. We use Poisson models, as there are a handful of CEOs who earn very low salaries in the pre-product period and Poisson regressions have better properties when handling outliers than OLS regressions after a logarithmic transformation (Silva and Tenreyro 2006). The regression coefficients in the table can be interpreted as percentage changes in the dependent variable with respect to a change in a regressor, with the Poisson model conditional mean function specified as

$$\log(\mathbb{E}(\text{Comp}_i | X_i, \text{Controls}_i)) = X_i\beta + \text{Firm Age}_i + \text{Controls}_i. \tag{1}$$

To account for the possibility that we may have repeated observations on the same firm (which we cannot observe because the survey waves are not linked over time), we cluster standard errors by founding year- region-industry cells.

The X matrix contains indicators for milestones, like achieving product status or raising capital, that may be realized over the firm lifecycle. The first column of Table 2 contains our baseline model with the fewest controls. Cash compensation is positively related to firm revenue, with substantial increases in compensation among firms that have positive revenue relative to the baseline category of pre-revenue firms. Column 2 adds firm age. After controlling for age, having positive revenue between zero and \$10 million annually is associated with an approximately 49% increase in pay ($\exp(0.401) - 1$) relative to pre-revenue firms. Compensation then further increases with revenue until firms hit more than \$100 million in annual revenue.

Column 3 tests for the importance of having a product for compensation, with striking results. This column includes the Post-Product Definition dummy for whether the firm has made it beyond the Pre-Product stage, as described in Table 1 and Figure 1. This separates the degree to which revenue and customer traction drives cash compensation or, instead, the existence of a tangible piece of intellectual property or technology that yields an increase in founder-CEO cash compensation. Even if the product is not yet sold, it may signal that the enterprise is on the path from differentiation to standardization, that the CEO is becoming a manager rather than an explorer or idea generator, and that the firm is separable from the founder (Rajan 2012). Having a product, rather than the transition to earning revenue or having paying customers, is the significant inflection point for compensation. Between Column 2, which does not have the Post-Product dummy and Column 3 which does, the coefficient on having between zero and \$10 million in revenue falls from 0.40 to -0.05. The 0.62 coefficient on the Post-Product dummy explains why. The compensation-revenue gradient remains positive for firms with over \$10 million in sales, but it is far less pronounced relative to Column 2. The product definition dummy, revenue dummies, and firm age together explain 48% of the variance in cash compensation. Subsequent columns add additional characteristics, as noted in the bottom of the table. The 0.365 coefficient after controlling for cumulative VC investment, total rounds of funding, industry, region, and firm age, implies that having a tangible product is associated with a cash compensation increase of approximately $\exp(0.365) - 1 = 44\%$. Together these additional controls raise the R-squared to 0.53, while product definition remains economically significant and meaningful. Region fixed effects are jointly insignificant in a Wald test in Column 5. Thus, considering firm-, industry- and region-level factors, the majority of the variation in pay is driven by variation in underlying *firm* fundamentals, swamping the geographic and industry differences that have been shown to be important in other contexts (Moretti 2010).

Column 6 addresses whether the results from the survey data generalize to the universe of startup firms by re-weighting the AHR sample to match the cross-section of VentureSource in each survey year. We target firm age and capital raised in this re-weighting exercise, as these dimensions capture potential sampling bias over the firm age distribution or a bias toward better capitalized firms in the AHR survey. The main results remain very similar to those without the re-weighting.

In Figure 2, we display the predicted densities of the level of cash compensation. We take the fitted values from Column 5 of Table 2 and show how the distribution changes for firms with different product statuses and revenue levels. Pre-product firms have a narrow density of pay, with a low mean. Post-product but pre-revenue firms have a distribution that is shifted to the right. The pay distribution continues to shift rightward as firms gain revenue.

2.1.5 Robustness of Compensation Results

In this section, we document the results of several tests to show the robustness of the finding that liquid cash compensation increases substantially after developing an initial product.

First, we probe whether our estimates are robust to unobserved firm or founder quality, and conclude that unobservables would need to be implausibly large to explain the increasing cash compensation that comes with having a product. Note that since survey waves cannot be linked over time, we do not observe the evolution of CEO compensation within a given firm. Instead, we exploit differences in the evolution of firms across their lifecyles in the cross-section of our surveys, tracing out the evolution of compensation based on firms observed at different points in their lifecycles. Because of this, one might be concerned that our estimates are contaminated by survival bias. For example, if only the highest quality firms from the pre-product distribution are observed among the post-product set, then the differences in compensation that we observe across firm age could be due to the survival of firms where the CEOs had higher pre-product salaries and the failure of firms that had lower pre-product salaries.

To assess how extreme this selection would need to be to change our conclusions, we conduct a bounding exercise in the spirit of Lee (2009). The intuition behind this exercise is to 'back out' how extreme selection would need to be for it to be the primary driver of a set of results. The exercise assumes that sample selection is monotone, in that firms with higher pre-product CEO compensation are more likely to be observed in subsequent rounds, which is likely to be the nature of selection in our case. In this case, one can compare the conditional mean for the top X% top of the distribution of pre-product firms' compensation with the distribution of post-product compensation, where X is the share of firms that survive. In other words, suppose only the top X% of firms in a given round survived to the next round. One can compare the distribution of compensation for the top X% of firms in the prior round to the distribution of compensation in the focal round. If the distribution of salaries is still right-shifted when one assumes such extreme selection, then selection cannot account for all the effect being measured. Appendix A.2 provides details about the procedure. Against the trimmed distribution that assumes surviving firms started as the most highly compensated among pre-product firms, we find an increase in log total compensation of 0.35 (a 42 percent increase in levels) due to having a tangible product (the standard error is .05). The mean (and median) of the post-product total compensation distribution are equivalent to the 91st percentile of the pre-product total compensation distribution. Given these extreme differences, selection on unobserved quality is unlikely to explain the importance of product milestones for CEO compensation in the cross-section of firms.

Second, we show that founder-CEO compensation in post-product firms appears similar in different eras and across industries. To examine the robustness of the high CEO-compensation over time, we bring to bear two different sources of data on publicly-traded companies, as shown in Figure 3. Panel (a) uses hand-collected data on corporate filings of new IPOs among VC-backed startups over the 2006–2018 period to document the total cash compensation for founder-CEOs of these newly public firms.[8] Every firm filing an S-1 for an IPO must provide 3 years of compensation history for the CEO and other top executives. Using these recorded filings, we examine cash compensation in 2015 dollars for firms that IPO over the time series and whose CEOs are still the founders.[9] Among firms that eventually go public, median compensation for founder-CEOs ranges between $290,000 and $400,000, which is consistent with the total compensation figures noted in Table 1. Compensation is at the higher end of this range for the latter years of the sample, which is consistent with firms taking longer to go public and therefore being somewhat larger. In Panel (b) of Figure 3 we draw on data from Execucomp and scraped proxy filings to plot bin-scatters of the relationship between CEOs compensation and the revenue of the firms they run. Consistent with the literature on CEO compensation, it shows that log CEO compensation scales linearly with the log of firm revenue for publicly traded firms. Two other features are worth noting: first, the slope of the relationship between CEO compensation and firm revenue is similar for both public companies and those in AHR suggesting that the 'manager contract' emerges early in the life of ventures once there is a product. Second, it can be seen that there is overlap between the compensation levels between the larger firms in AHR and the smaller publicly-traded firms. Together, these patterns provide reassurance that the AHR

[8]Baker and Gompers (1999) collected similar data for their analysis of the transition of VC-backed startups from private to public firms.

[9]If the S-1 does not contain compensation data, then we collect the first post-IPO DEF 14-A proxy statement filing.

data is capturing systematic moments of the late-stage, post-product compensation contract rather than either a shift over time or a blip driven by high inflows of capital into the VC sector during the low interest rate environment.

An additional dataset from Compensation Pro (formerly provided by Dow Jones), covering 2008–2010, provides suggestive evidence that our post-product compensation moments for tech firms extend to other industries, including industries that are not available in our AHR data. Bengtsson and Hand (2011) have used this data to understand how compensation changes with fundraising events, and they document funding milestones that increase compensation. Unfortunately, these data are not well suited to capture pre-product firms and the inflection point around the product-transition. As we discuss in the note to Appendix Table A3, the share of pre-product startups in the Compensation Pro data is 2.5%, whereas 54% of VentureSource startups alive between 2008 and 2010 are classified as pre-product. Lack of representativeness among early stage firms means that we use the data to assess the similarity of compensation for *post-product* firms from this era.

The post-product compensation moments we observe in the AHR data for tech companies in 2015 and 2017 appear similar to the compensation moments for post-product tech companies and non-tech companies (which are not observed in AHR) in the 2008–2010 Compensation Pro data. Table A3 displays the means and medians across different industries. For IT firms in the Compensation Pro data, the median compensation for a post-product founder-CEO is over $225,000. The median is exactly $200,000 in the AHR data. For other industries, it is clear that median compensation for post-product CEOs is in a similar range. All industries have medians that are at least $185,000 and the corresponding averages are somewhat higher. These moments suggest that the compensation patterns from the AHR data are applicable to other industries and time periods as well.

2.2 Founder-CEO Replacement

Our next set of results looks at founder turnover before and after achieving product status. Rajan (2012) theorized that at the birth of a new firm, the key human capital in a venture is not replaceable, requiring the entrepreneur to have significant control to create a differentiated idea. This key human capital becomes more replaceable once the firm undergoes a transition from differentiation to standardization.

If indeed an initial product is a fundamental milestone associated with the transition from differentiation to standardization, we should expect to see very few instances prior to a firm achieving product status where the founder is replaced but the firm survives. In other words, founder exit is likely to be synonymous with firm exit pre-product. However, the hazard for founder-exit and firm-exit should diverge after a firm has a product.

Consistent with this view, the summary statistics in Table 3 reflect an increased likelihood of observing non-founders in the CEO role in the AHR data after firms have a product, but not before. Only 2 percent of pre-product firms are lead by a non-founder CEO. This number jumps to 9% for post-product, pre-revenue firms and then increases as firms gain revenue.[10] This table also provides summary statistics about the components of non-founder CEO compensation relative to founders. While non-founder CEOs are paid more in levels, the gradient in pay with respect to firm revenue is similar for founders and non-founders; founders also hold more equity in the firm.

Table 4 shows that after controlling for age, industry, and geography, founder-CEO turnover increases after the firm has a product. The table reports results from estimations where an indicator for having a non-founder in the CEO position is regressed on a post-product indicator. We conduct this analysis separately on the AHR data and using VentureSource. The first three columns report regressions on the AHR data of a non-founder on the Post-Product definition dummy along with an increasing number of controls. The coefficient on the Post-Product definition dummy lies between 0.044 to 0.052. This analysis conditions on surviving firms, suggesting that firm failure rather than leadership turnover, occurs for pre-product firms that are unable to achieve product development milestones.

To investigate whether the connection between turnover and product holds in a larger sample, Columns 4 and 5 report similar regressions using the VentureSource sample for the cross section of firms in 2015 and 2017 that were backed by VCs firms participating in the AHR surveys. Because VentureSource provides a view of the executive team at every financing event, we can accurately assign turnover events in an interval around financings. In this sample, 1.2% of pre-product firms are lead by a non-founder as CEO. The post-

[10] Prior studies about why founding CEOs are replaced point to bi-modal reasons for turnover (Wasserman 2003). Some turnover occurs in firms that are struggling (Ewens and Marx 2017). Other firms experience turnover when venture investors perceive the need for extremely fast growth for which founders are ill-equipped. The canonical example is Google, where Eric Schmidt was brought in to provide "adult supervision."

product dummy has a coefficient of around .04, implying again that turnover is substantially more likely post-product than pre-product for these firms. Finally, Column 6 uses a firm-year panel up to the year of founder-CEO replacement, allowing us to fit a hazard model of founder-CEO turnover. The sample includes all startups first financed on or after 2000. Because the data is set up in an event history format, the mean of the dependent variable is not directly comparable to prior columns. The hazard of having a founder-CEO being replaced is 0.5% per year prior to achieving product. Again, as seen in Column 6, having a product increases the replacement hazard significantly. The 0.019 coefficient on Post Product Definition, combined with the constant, indicates that every year after achieving product status, the CEO replacement probability is about 2.4 percent, which is over 4 times greater than the CEO replacement hazard for pre-product firms.[11]

3 Founder CEO Compensation, Time to Product, and the Attractiveness of Entrepreneurship

Thus far, we have documented the importance of a startup's initial product as a key milestone in its lifecycle. At the extensive margin, this marks the point when startups are decoupled from their founders and can continue to survive under new CEO management. At the intensive margin, cash compensation for CEOs rises after this milestone, even for founder-CEOs who were previously paid a fixed, low salary. A natural question that arises then is whether the increase in cash compensation for founder-CEOs post product is sufficient to insure them from the non-diversifiable risk they face at the birth of new ventures. If so, this can impact the attractiveness of selection into entrepreneurship. We address this question in this Section.

To understand which potential founders we would expect to be willing to enter entrepreneurship and which would not, we build on one of the best-known benchmark models in this literature, by Hall and Woodward (2010). Hall and Woodward model a consumption-savings problem that incorporates many realistic features of the payoffs to founders, includ-

[11] Note that the hazard estimates are not directly comparable to the cross-sectional point estimates in Columns 4 and 5 of Table 4 because the sample includes all years up to a founder-CEO replacement event, whereas the cross-sectional regressions only look at the frequency of observing a replacement at a single snapshot in time.

ing subordinate positions as common equity holders. A forward-looking founder understands that venture investors have liquidation preferences, while exit values are stochastic and the time to a liquidity event is unknown ex-ante. They are also among the first to use realistic risk preferences while modeling entrepreneurs' consumption and asset accumulation decisions. Their baseline model can be used to assess the attractiveness of VC-backed entrepreneurship for potential founders with different combinations of outside salary and wealth.

Hall and Woodward (2010) approximate founder-CEO's compensation from birth to an exit with a flat salary of $150,000. However, as documented above, a relatively low and flat salary contract is only applicable pre-product, while compensation rises substantially after a startup achieves a product. Understanding the degree of non-diversifiable risk founders face therefore requires that the entrepreneur's value function in the consumption-savings problem is split into a pre-product and a post-product phase, with a transition probability between these phases that determines the amount of time an entrepreneur expects to be holding the low, flat salary associated with the pre-product 'entrepreneur contract.' We build a model that incorporates these features, but we begin with a baseline model to build intuition.

3.1 Baseline Model

In the baseline model, the key consideration faced by a risk-averse potential entrepreneur is to understand whether the consumption utility if they pursue an entrepreneurial idea is sufficiently large relative to their current outside option (their salary in wage employment) so as to attempt this risky activity. Risk arises because the firm value that accrues to an entrepreneur in a liquidity event is a random variable with substantial variability. Prior to a liquidity event, these future assets do little for consumption, as Hall and Woodward make the realistic assumption that there is no borrowing against equity. As a result, the entrepreneur must consume out of either salary or wealth generated prior to founding. To fix ideas, let π_{t+1} be the probability of a liquidity event at date $t+1$ with the distribution of the stochastic after-tax payoff to the entrepreneur, X_t, conditional on an exit at $t+1$. We write the ex-ante value function for an entrepreneur with assets A_t at time t as:

$$
\begin{aligned}
V(A_t) &= \max_{c_t<A_t} u(c_t) + \frac{1}{1+r}(1-\pi_{t+1})E_{w_t}V((A_t-c_t)(1+r)+w_t) \\
&+\frac{1}{1+r}\pi_{t+1}E_{w_t,X_{t+1}}V((A_t-c_t)(1+r)+w_t+X_{t+1})
\end{aligned}
$$

The entrepreneur solves for a consumption path with expected entrepreneurial wages while knowing the post-exit value function takes the form

$$
V^{AfterExit} = \frac{1+r}{r}u(\frac{rA+w^*}{1+r}), \tag{2}
$$

with w^* the non-entrepreneurial wage. Hall and Woodward further assume that the non-entrepreneurial wage is the same before and after founding, which appears to be a reasonable assumption based on the difference between pre- and post-founding average occupational earnings that we present in Appendix B.3 and Appendix Figure B1.

Hall and Woodward assume that the flow utility is isoelastic, with $u(c_t) = \frac{c^{1-\gamma}-1}{1-\gamma}$, where γ is the coefficient of relative risk aversion. A somewhat standard level of risk aversion sets $\gamma \approx 2$. They also assume that entrepreneurs earn an annual pre-tax salary of \$150,000 over the entire life of the firm and consume or save out of post-tax earnings. Their baseline model provides wealth thresholds such that the certainty-equivalent value of entrepreneurship is positive at different levels of the non-entrepreneurial wage, w^*. When the certainty-equivalent is positive, founding has positive expected utility and entry is incentive compatible.

3.2 A Model with Pre- and Post-Product Stages

We extend the Hall and Woodward model in three important ways. First, we allow w_t to be firm age and product-status dependent. Second, we show that the value function can be split into pre-product and post-product phases, with a transition process between the phases. We account for the transition to product status or exit for pre-product firms and estimate the transition hazards from the data. That is, pre-product firms either transition to post-product status, exit prior to achieving a product, or continue product development. The timing of these transitions relative to the duration of the product development period potentially changes the risk founders bear. Finally, our setup allows us to examine comparative statics, like taking a year longer to achieve a product, or taking a year longer to exit as a post-product firm. We relate these comparative statics to the attractiveness of founding across the income distribution, providing testable implications of the model.

Including these extensions requires additional notation. We denote by p_{t+1} the probability that a pre-product firm at age t will transition to having a product by $t+1$. Exit probabilities depend on product status and firm age, with $\pi^{\sim P}_{t+1}$ and π^{P}_{t+1} the respective probabilities of exiting at age $t+1$ without and with a product. We then denote by $E_{X_{t+1}|\sim P}$ and $E_{X_{t+1}|P}$ the expectation of the founder's share of the exit value conditional on exiting at $t+1$ for each product status.

The value function for post-product firms is little changed from the baseline setup. It is

$$\begin{aligned} V(A_t)^{PostProduct} &= \max_{c_t<A_t} u(c_t) + \frac{1}{1+r}(1-\pi^{P}_{t+1})E_{w_t|P}V((A_t-c_t)(1+r)+w_t) \\ &\quad + \frac{1}{1+r}\pi^{P}_{t+1}E_{w_t,X_{t+1}|P}V((A_t-c_t)(1+r)+w_t+X_{t+1}). \end{aligned}$$

With probability π^{P}_{t+1} the firm exits and the founder realizes a stochastic payoff associated with an exit from the post-product distribution of exits at age $t+1$. Then in the next period, the after-exit value function applies from equation (2), and the founder has any assets that remain from the period when the startup was active plus the assets that come from an exit. With probability $(1-\pi^{P}_{t+1})$ the firm continues and the entrepreneur receives a wage realization from the firm-age specific distribution of wages conditional on having a product.

The value function for pre-product firms is

$$\begin{aligned} V(A_t)^{PreProduct} &= \max_{c_t<A_t} u(c_t) + \frac{1}{1+r}(1-\pi^{\sim P}_{t+1}-p_{t+1})E_{w_t|\sim P}V((A_t-c_t)(1+r)+w_t) \\ &\quad + \frac{1}{1+r}p_{t+1}V^{PostProduct} \\ &\quad + \frac{1}{1+r}\pi^{\sim P}_{t+1}E_{w_t,X_{t+1}|\sim P}V((A_t-c_t)(1+r)+w_t+X_{t+1}). \end{aligned} \tag{3}$$

Without a product, there are three state transitions. With probability p_{t+1} the firm transitions to the post-product value function. With probability $\pi^{\sim P}_{t+1}$ the firm exits without a product. This means with probability $(1-\pi^{\sim P}_{t+1}-p_{t+1})$ the firm continues with the pre-product value function.[12]

3.3 Model Calibration

3.3.1 Data Inputs for Calibration

We use two main sources of data to calibrate the model. Data related to firm-level outcomes and transitions are drawn from VentureSource. The projected evolution of founder compensation is based on data from AHR and is supplemented by Compensation Pro moments from earlier periods and different industries.

The VentureSource data contains information on the firm's industry, total exit value for firms that IPO or that have been acquired, the firm's founding year and age at exit, the firm's total capital raised, and the firm's product transition status. From the raw data, we compute product transition-by-firm age probabilities, exit probabilities-by-age conditional on product status, and the conditional distribution of exit values to the entrepreneur given product status for an exit at $t+1$. Like Hall and Woodward (2010), there are a number of assumptions and imputation steps that we use to make the data suitable for calibration. These include imputing missing time to exit/failure for so called zombie firms, imputing missing acquisition values, adjusting for unobserved timing of product transitions between funding rounds, and calculating the entrepreneurs' share of firm value in the presence of liquidation preferences.[13] Appendix B.1 and Appendix Table B2 detail how we deal with these issues and the number of observations that are affected by imputations.

In choosing the time period over which we use firm-level data to calibrate the model, we attempt to balance a tradeoff: on the one hand, we are interested in a long time series that

[12] Once the firm transitions to the post-product value function, the exit probability and X do not depend on the age when the product transition occurs. This assumption appears to be a reasonable, as the average firm-level exit value in VentureSource data does not depend on the age when the firm achieves product status.

[13] Calculation of the entrepreneur's share of each total exit requires assumptions because the data do not contain investors' liquidation preferences or the individual founder's equity share. As detailed in Appendix B.1.2, we assume that investors exercise their liquidation preferences based on thresholds of the total exit value relative to capital invested. Gornall and Strebulaev (2020) show the importance of different share classes and liquidation preferences for valuing the equity in privately held companies, and our procedure recognizes that founders' common equity holdings are not a sufficient statistic for payoffs, even when the exit value of the firm is known, due to the subordinate position of common equity. For individual founder common equity holdings, we use average founders' shares of common equity from AHR data by age and product status.

covers multiple venture capital investment cycles (Gompers and Lerner 2004). On the other, we need to be careful not to truncate outcomes since we are interested in understanding the joint distribution of exit values with firm age. This is of particular importance given the documented increase in time to exit for some of the most successful firms (Ewens and Farre-Mensa 2020). We balance this tradeoff by studying exits anytime between 1987 to 2019, but we restrict the main sample to firms founded between between 1987 and 2006 so as to enable sufficient time to follow even the most successful firms to exit. We track firms' outcomes through the end of 2019 and convert all exit values to 2006 dollars using the CPI for all urban consumers.

We use the AHR compensation data to specify the stochastic process for the evolution of founders' future wages and cash compensation. These data enter the model through $E(w_t|P)$ and $E(w_t| \sim P)$. Two possibilities for specifying these expectations are: a) use the raw realizations of total compensation in the data conditional on firm age and product status or b) use moments of the wage given age and product status to form expectations, but build in negative shocks that approximate financial distress. While both approaches yield similar results, we have opted to feature the moment-based approach because future researchers will be able to implement this procedure without having access to the compensation data. We set $E(w_t|P) = .9\overline{w_{t,Product}} + .1 \times w_{Low}$ where $\overline{w_{t,Product}}$ is the average cash compensation observed for a firm aged t with a product and w_{Low} = \$120,000. This expectation is a convex combination of the observed mean of cash compensation given age and product status in the AHR data and a low draw, which reflects the potential for future firm distress that reduces salary. For firms without a product, we set $E(w_t| \sim P) = \overline{w_{t,\sim Product}}$ from the AHR data. Comparing the two expectations, the post-product wage process has a jump at product status, grows with firm age, and (because the agent is risk averse), features a probabilistic salary cut that reverts nearly back to pre-product levels, w_{Low}. Because w_t is post-tax income, we compute average tax rates for different values of pre-tax income using the NBER's Statistics of Income database of sample tax filings.

From the model, we solve for the pre-entry salary and accumulated wealth combinations for which entrepreneurship is attractive given the product transition process, evolution of salary, time to exit, and distributions of exit values. For this calculation, we find the minimum value of wealth for a given outside salary w^* such that the pre-product value function in (3) exceeds the value function from not entering entrepreneurship, which we characterize for each wealth-salary combination using equation (2). We solve the model numerically. For details see Appendix B.2.

3.3.2 Baseline Results

Figure 4 presents the first key output of our calibration. Each line in the figure traces out the minimum wealth required for founding to be attractive under a given outside salary. The area above each line represents the region where founding a VC-backed startup is incentive compatible. The region below each line represents missing entrepreneurs. The top line in the figure represents salary-wealth thresholds for entry to be attractive under the compensation contract that pays the founder a fixed salary of $150,000 over the life of the venture. The middle line, labeled "AHR," represents the compensation contract that changes with product status and firm age. Finally, the lower line illustrates what would happen to the attractiveness of founding using a thought experiment where an entrepreneur can bypass the pre-product phase altogether and directly establish a post-product company.

The first insight from the figure is that potential entrepreneurs who earn less than $150,000 in pre-entry salary are not constrained by the burden of non-diversifiable risk from founding. The point furthest to the left on the X-axis represents an entrepreneur with an outside salary of $150,000. The wealth requirement for those earning $150,000 or less is $0, as the salary in a new venture would meet or exceed their outside earnings ability. Anyone earning less than $150,000 in the outside labor market would be able to at least maintain their level of pre-founding consumption upon entering entrepreneurship.

Second, the figure clearly illustrates that non-diversifiable risk is most likely to deter entrepreneurial entry from those with high salaries and limited wealth. Under the fixed pay contract, a potential entrepreneur earning $400,000 in the outside labor market would require $1.2 million in liquid wealth to enter. An entrepreneur with a $500,000 outside salary would require $2.5 million in wealth, while a $600,000 outside salary would necessitate $4.5 million in wealth and a $900,000 outside salary would require nearly $17.8 million in wealth. Under the flat salary contract, as the potential founders' outside salary increases, the opportunity cost of entering entrepreneurship increases, which requires consumption smoothing out of savings. Wealthier individuals are able to finance this consumption in light of the possibility that the equity value upon exit will lead to future asset accumulation.

Third, compensation that responds to product milestones makes founding more attractive for high earning individuals relative to the flat salary of $150,000. The wealth threshold for founding to be incentive compatible falls from $1.2 million to $640,000 for someone with a $400,000 outside salary and from $2.5 million to $1.5 million for a $500,000 outside salary. Despite average pay that is lower than $150,000 in the pre-product period, the product signal has value when viewed through an experimentation lens (Kerr, Nanda, and Rhodes-Kropf 2014; Manso 2016; Dillon and Stanton 2018), as CEOs who do not achieve product status can fail quickly and do not persist in the venture. Above $500,000 in pre-entry salary, the simple change in the compensation contract may not be enough to induce many higher earners to enter. Although the wealth threshold falls from $4.5 million to $2.9 million for someone with an outside salary of $600,000, the $2.9 million in wealth is 7.9 times the annual post-tax pay for this potential entrepreneur. Someone with a $600,000 outside salary would need to be an extreme saver or hold external wealth to justify entrepreneurial entry.

Fourth, faster time to product can play a substantial role in reducing the burden of non-diversifiable risk for those at the top of the earnings distribution. Eliminating the pre-product phase entirely reduces the wealth threshold for someone with a $600,000 salary to $1.5 million, which is roughly four times post-tax salary. As this model holds fixed time to exit and the distribution of exit values, it demonstrates that time to product is an important determinant of the risk borne by entrepreneurs. Faster time to product unlocks greater cash compensation earlier in the life of a startup, alleviating some of the burden of non-diversifiable risk. Of course some risk remains for the very highest earners, but the dynamic compensation contract after product milestones greatly narrows the earnings gap for those with high incomes.

Fifth, the figure highlights the importance of considering the appropriate benchmark when considering potential "missing entrepreneurship." On the one hand, it shows that despite non-diversifiable risk potentially deterring top earners from entry, most of the population has outside salaries such that entrepreneurship looks attractive. To show this, we note the population percentile of the income distribution corresponding to outside compensation below the x-axis in Figure 4.[14] The right y-axis contains percentiles of the wealth distribu-

[14]These estimates are taken from the NBER's version of the IRS Statistics of Income files and utilize data from W2 filings for a stratified random sample of tax returns. We use data items 85 and 86 from the 2001–2016 SOI data, which gives W2 earnings for individual filers and married joint filers. Individuals without W2 earnings are not included in the percentile estimates.

tion such that entrepreneurship has a positive certainty equivalent.[15] Ninety-five percent of the population earns under $150,000 and is not constrained by non-diversifiable risk.

The incentive compatibility question only begins to become relevant at income levels exceeding around $225,000, which is the top 2.1% of the population with taxable labor income. Although we do not observe the joint distribution of income and wealth, we can approximate how many high-earners are deterred by non-diversifiable risk by treating the income and wealth distribution as bivariate lognormal, with a correlation coefficient taken from the literature.[16] Using this approach, about 469,000 high-earners (or 14% of the population of high-earners) would be deterred from founding under the product- and age-dependent contract.

We note that the missing individuals under this exercise, while a relative small share of the workforce, are likely those who are extremely consequential for innovation. That is, missing entrepreneurs are high human capital individuals, as reflected by their high salaries (that correlate with productivity). In other words, once the risk set is adjusted to account for the higher income of those with the highest levels of human capital, the potential consequences of missing entrepreneurs can be substantial, particularly when considered in light of the fact that around 1500 VC-backed startups are founded every year.

3.3.3 An Extension to Startup Joiners

Our main focus is on founder-CEOs of startups, as they are the source of the 'differentiated' ideas. However, our approach can also apply to startup employees or joiners, who are potentially important inputs for innovative companies. Our data use agreement with AHR precludes us from disclosing compensation moments for non-senior (e.g. C-level, non-CEO) employees, but we can analyze the attractiveness of joining a startup at the senior

[15] These wealth percentiles are interpolated from data provided in Saez and Zucman (2016). To perform this interpolation, we take the log of the minimum wealth at different percentiles of the distribution and the log of the percentile and then use linear interpolation. We then exponentiate the interpolated log percentile.

[16] To recover the number of missing entrepreneurs, we first estimate moments of the marginal distributions of log income and log wealth using the quantiles that we report in Figure 4. We assume that the correlation between income and wealth in levels is 0.48, which we take from Díaz-Giménez, Glover, and Ríos-Rull (2011). We then generate bivariate normal random variables with covariance $\log(0.48\sqrt{(exp(\sigma_I^2) - 1)(exp(\sigma_W^2) - 1)} + 1)$ and we calculate the share of draws with log income above $225,000 where a potential entrepreneur would be deterred from founding based on the income-wealth thresholds from the model. We multiply this share by a labor force of 150 million people.

level. We use cash and equity compensation data for these individuals as inputs into the consumption/savings models that we use to study the attractiveness of entry for founders. Appendix B.4 presents these results. The main finding is that very few high-income individuals will join startups as non-founders at the senior level. Founders' equity stakes and cash compensation are crucial to overcome the burden of non-diversifiable risk, but joiners' are paid less than founders, on average, and their equity stakes are lower. These differences make joining a startup as a non-founder harder to justify for someone with attractive outside options. Those with the high outside earnings potential would need strong non-pecuniary motivations to want to join a VC-backed startup at its earliest pre-product stage.

3.4 Quantifying Importance of Time to Product

The discussion above has highlighted the importance of expected time to product in driving selection into VC-backed entrepreneurship. We turn next to quantifying its importance using variation in the observed time to product across industries and time. We believe this is a fruitful exercise for two reasons. First, our extended model is well-suited to understand comparative statics with respect to time to product. We can use it, for example, to understand how changes in time to product across industries compares with other documented changes, like firms staying private longer prior to an acquisition or IPO. With extended times to acquisition/IPO, the founder-CEO goes longer without access to the illiquid wealth in the venture; but increases in cash compensation that arrive sooner due to faster time to product may partially or fully offset the impact on founder risk-bearing from firms that remain private longer. We can study how individuals with different salary-wealth profiles evaluate the entry decision based on the above trade-offs.

Second, while credible assessments of the number or cost of 'missing entrepreneurs' in the aggregate are difficult, a comparison across industries can isolate the degree to which industry-level variation in time to product impacts the amount and composition of entrepreneurial talent entering these sectors. Industry and time series variation in time to product enables us to assess relative differences in missing entrepreneurs, providing some of the first evidence on the particular sectors where the supply of innovation talent is most constrained. We address each of these questions in turn.

3.4.1 Industry-level Variation in Time to Product

To isolate the importance of the product transition process, we exploit the introduction of cloud computing by Amazon Web Services in 2006, which unexpectedly and differentially impacted the time to product over the subsequent years. As described in detail by Ewens, Nanda, and Rhodes-Kropf (2018), who quote Amazon's Chief Technology Officer at the time, the capability behind the cloud computing platform was initially developed to service internal operations, and only later seen as possible source of revenue by servicing outsiders. Cloud computing allowed internet-oriented startups to rent (rather than previously buy) expensive equipment for their initial operations – which lowered the cost and time associated with initial experimentation through faster prototyping, reduced need for firm-specific capital investment – and thereby sped time to product. Other industries, such as those in consumer products, healthcare or energy and industrials, were less affected by cloud computing. Ewens, Nanda, and Rhodes-Kropf (2018) identify 2006 as the year when Amazon began to make cloud computing products available to users, so we use 2006 as the marker for the end of the pre-period in our data. Firms founded after 2007 are in the post-period.

Specifically, we split the data by cloud and non-cloud affected industries as defined by Ewens, Nanda, and Rhodes-Kropf (2018). We separately compute product transition rates and exit probabilities for cloud affected and non-cloud affected industries in the pre-period and the post-period. Our goal is to understand how observed changes by industry and over time in the speed to product impact the attractiveness of entrepreneurial entry. Note that although we compute separate product transition rates for each of cloud and non-cloud affected industries in both the pre- and post period, we hold fixed the exit value distribution for each of the cloud and non-cloud affected industries, since as noted earlier, exit values are not observed for many firms founded during the 2007–2014 period.[17]

As expected based on Ewens, Nanda, and Rhodes-Kropf (2018), the firms in cloud affected industries transitioned to having a product (or failure) more quickly after 2006. Among the startups founded in cloud-affected industries between 2007 and 2014, 29.7% remained pre-product three years after founding compared to 34.9% that remained pre-product in the pre-period, a reduction of 15%. Although firms in non-cloud affected industries took longer to achieve product status at baseline (56.1% remained pre-product by year three in

[17] It seems reasonable that potential entrepreneurs are likely to use the historical distribution of exit values to inform beliefs about the attractiveness of founding.

the pre-period), there was minimal change in the rate of product-transition by year 3 in the 2007-2014 cohorts of new firms in these industries (to 55.6%). This enables us to calibrate a theoretical difference-in-differences model that illustrates how the attractiveness of founder entry changes across industries and eras. Appendix B.1 provides the estimates of p_{t+1} and $\pi^{\sim P}_{t+1}$ for the pre- and post-periods across cloud-affected and non-cloud affected industries. Our comparative statics analysis uses these objects to study how the attractiveness of founding changes under different times to product. We also consider another possible change over time, which is longer times to IPO, as firms stay private longer.

3.4.2 Results Across Industries and Time

We present differences in industry attractiveness and comparative static results in Table 5. Panel A presents wealth thresholds given an outside salary for firms in cloud affected industries and Panel B provides the same thresholds for industries that were not affected (or were less affected) by cloud computing. The first column presents salary-wealth thresholds in the pre-period assuming that founder-CEOs earn a flat salary of $150,000 per year. The across-industry comparison of the flat salary contract shows that cloud-affected industries were more attractive even absent a product-dependent contract. For example, in cloud-affected industries, it would have taken at least $3.9 million in wealth for someone with a $600,000 outside salary to find entrepreneurship attractive, while the threshold rises to $6.3 million for non-cloud affected industries.

The product- and age-specific pay contract (Column 2) makes cloud-affected industries even more attractive for high-earners in the pre-period because of faster time to product. With the product-dependent compensation contract, the wealth required for someone with a $600,000 ($750,000) outside salary to found in a cloud-affected industry falls from $3.9 million to $2.4 million ($8.4 million to $5.6 million). These thresholds are lower than those presented for the overall sample, in Figure 4 as firms in cloud-industries in the pre-period had relative quick product development cycles compared to those in life sciences, industrial, and consumer product verticals. The equivalent thresholds change from $6.3 million to $4.3 million ($13.4 million to $10.5 million) for non-cloud industries. Using our parametric joint distribution approach to compute missing high-earning founders, we present changes in the number of missing entrepreneurs relative to the flat pay contract in Column (1). The product- and age-dependent contract makes founding financially attractive for approximately an additional

252,000 high-earners in cloud-affected industries and 170,000 additional founders in non-cloud industries.

Figure 5 provides the graphical comparison between industries that underlie these results.[18]

The bottom line of Figure 5 and Table 5 Column 3 show that cloud affected industries' more rapid time to product resolution in the post-period increases the attractiveness of founding, especially for those with relatively high outside salaries. Column 3 holds fixed the time to exit and the distribution of exit values from the pre-period, while applying the post-period product transition processes. Wealth thresholds fall in cloud affected industries, drawing in an additional 313,000 potential founders relative to the flat pay contract in the pre-period and an additional 61,000 relative to the pre-period contract that uses the AHR pay moments. Recall that in the pre-period, the model suggests that non-diversifiable risk potentially deterred 14% of individuals in the top 2% of the United States earnings distribution from founding. The 61,000 new potential founders in software and related sectors after the introduction of cloud computing represents about 13 percent of those deterred from founding by non-diversifiable risk and about 2 percent of the total population earning more than $225,000.

For non-cloud industries, the modest changes in product time draw in only about 24,000 additional founders relative to the pre-period AHR contract. Faster time to product, combined with compensation that responds to product milestones, makes entry more attractive.[19]

Finally, the last column of Table 5 allows us to compare results when we increase the time to exit simultaneously with the faster time to product. When comparing these results

[18] While our main analysis uses the industry-specific distribution of exit values, Appendix Figure B2 shows that founding in cloud-affected industries appears more attractive when we hold fixed the exit distribution between cloud and non-cloud industries. That is, when we fit the model on the overall exit value distribution but apply the different product transition rates, we find cloud affected industries look more attractive due to faster time to product. For a potential founder with a $600,000 ($750,000) salary, using a common exit distribution gives a wealth threshold of $2.5 million ($5.8 million) for cloud-affected industries and $3.6 million ($9.2 million) for non-cloud affected industries in the pre-period.

[19] It may be surprising that the change in high-earning founders in cloud affected industries relative to non-cloud industries isn't even larger, as the change in speed to product in non-cloud industries is small. Note that most of the founders drawn into non-cloud industries earn between $225,000-$300,000, which is a dense part of the distribution of high-earners. For these individuals, small changes matter. The change in the number of potential founders in cloud affected industries is more muted in this range despite the much greater changes in speed to product because many potential cloud affected entrepreneurs were already unconstrained in the pre-period due to the faster time to product at baseline.

to the pre-2006 era, we can disentangle the relative importance of faster time to product or longer time to exit (for successful exits). We find that the faster time to product dominates the longer time to exit in cloud-affected industries, as wealth thresholds are lower in Column 4 than in Column 2.

Together, these results have testable empirical implications. They predict that we should see a differential increase of founders entering VC-funded entrepreneurship from high-earning occupations in cloud-affected industries after 2006. For non-cloud industries, the modest changes in product timing draw in far fewer potential high-earning founders. While our analysis of the number of missing entrepreneurs may be sensitive to distributional assumptions, the approach is the same for cloud and non-cloud industries. This suggest that the relative change in the number of potential entrepreneurs drawn into founding is about twice as large in cloud affected industries compared to non-cloud industries.

4 Selection into VC-backed Entrepreneurship based on Time to Product: Empirical Analysis

Thus far, we have shown theoretically that the expected time to achieving a product is a key driver of the non-diversifiable risk faced by potential entrepreneurs. This is because developing an initial product marks the resolution of important uncertainty, enabling venture capital investors to shift founder-CEOs from an 'entrepreneur' to a 'manager' compensation contract. Moreover, our model calibration shows that variation in time to product across industries is quantitatively important for the attractiveness of entrepreneurship for individuals with high levels of outside earnings.

In the third step of our analysis, we conduct an empirical investigation of selection into VC-backed entrepreneurship. Our model, calibrated using firm-level variables and startup CEO compensation, makes predictions about systematic differences in the number of founders with high pre-entry earnings who select into industries with different times to product. The empirical analysis that follows allows us to directly test these predictions using entirely different data on the career histories of entrepreneurs.

The key testable prediction related to the background of founders stems from the fact that non-diversifiable risk is most likely to deter entrepreneurial entry for those with high salaries and limited wealth. While we do not observe wealth, we use LinkedIn data to determine the age and career histories of those who become founders, enabling us to estimate their pre-entry earnings by matching based on industry, occupation code and age to the US Census Bureau's American Community Survey. We turn next to describing the data used for the empirical analysis.

4.1 LinkedIn Data

We acquired a full download of public LinkedIn profiles as of 2017 from the data vendor DataHut. The data provider legally scraped (per a Supreme Court decision) individual and company profiles to create structured data on individual education and work experience. In total, the database includes 3,250,662 U.S.-based firms and 55,446,080 unique employees of such firms.

We form our analysis dataset by identifying founder-CEOs of US-headquartered VC-backed firms that were founded between 2001 and 2016.[20] In total, we were able to identify founder-CEOs of 9,574 firms with biographical data available in LinkedIn, corresponding to 36% of recorded startups over this period in the VentureSource Universe. Details of the procedure used to identify founder-CEOs, validation of match rates to VentureSource, and an analysis of observable differences in the startups where we did and did not find a match are available in Appendix C. Appendix C reports on bias in the LinkedIn sample relative to the VentureSource universe. While we find that the LinkedIn sample is selected, we do not detect differences in our LinkedIn coverage that correlate with the diffusion of cloud computing across industries. If anything, LinkedIn is more likely to capture stronger startups in the pre-period for both cloud and non-cloud industries, which should make it less likely for us to find an increase in high-earner founders in the post-period.

Having identified Founder-CEOs in LinkedIn, we use LinkedIn profile information to develop an understanding of their pre-founding education and career history. We identify undergraduate college education (university, degree, major, and graduation year) and any graduate degrees. When available, college graduation dates allow us to impute the founder's birth year. We observe approximate founder age from their educational records for 54%

[20] The start date of the sample is driven by the fact that LinkedIn data is extremely patchy prior to 2001. The end date is based on the fact that our LinkedIn sample was collected as of 2017 (i.e., the company selling the data ran its last scrape in early 2017).

of the sample. Later, when matching to the Census Bureau's ACS data, we approximate age for 33% of the sample with years of reported prior work experience if dates of school attendance are missing.

For work experience, the data provides an individual's title, start date, and end date (if applicable) at each of their employers. Company profiles include headquarter location, short business description, and firm-assigned industry classification. In some cases start dates and end dates are missing and thus we cannot use the founder-CEO's biographical data in the analysis below. Similarly, not all individuals in this sample have pre-founding experience or listed education. In both cases, for the purposes of reporting summary statistics we set related variables to missing (rather than zero) because LinkedIn does not require that a user complete their profiles, so missing data is not necessarily a signal of zero education. Our results on founder background and entry patterns in subsequent analyses are not sensitive to how we code founder backgrounds with missing information, including whether we ignore founders with missing data, use subsets of founder information that we do observe, or use imputed values.

Table 6 presents summary statistics for the sample of founder-CEOs we were able to identify in LinkedIn. On average, founder-CEOs are 35 years old and have over 11 years of work experience across 4 different jobs prior to their venture. They are also highly educated: about three quarters have a bachelors degree or higher, and more than half of them have education beyond a bachelors degree. The fraction with advanced degrees is significantly higher than the U.S. population. For example, the 13% of founders with a Ph.D. compares to 4.5% as reported by the U.S. Census Bureau's Current Population Survey Annual Social and Economic Supplement.

Looking across columns of Table 6, founder-CEOs of startups in industries not impacted by cloud computing appear somewhat older at founding, accounted for by their higher level of education and slightly longer work experience, on average. At first glance, the fact that these individuals founding startups in non-cloud affected industries are more educated and have more work experience might suggest they have higher pre-entry earnings. However, it is hard to make such an inference without an understanding of the specific job the individual held. For example, they may be systematically more likely to have come from occupations or industries that are lower paying despite having high levels of education. This ambiguity also highlights that relying solely on matching to LinkedIn has some limitations. We do not observe prior income, so we cannot directly validate our predictions around entry patterns among high-earners. This exploration requires a way to understand pre-founding earnings. Absent administrative data (where due to the small number of VC-backed founders, disclosure rules would make it difficult to report exact earnings), the best approach is likely to understand the conditional distribution of earnings, given occupation, industry, and age for founders compared to non-founders.

4.2 Match to Census Bureau's American Community Survey

We overcome these limitations by matching the LinkedIn founder-CEO sample with the Census Bureau's American Community Survey (ACS). The ACS data is an annual representative sample containing detailed data on 1% of US households. The ACS contains information on adults' education and labor force status. For those employed, it reports industry, occupation, age, and total labor income, enabling us to estimate the probability (from the ACS universe) that an individual founder had a high-earning position prior to entering entrepreneurship.

We use a matching on observables approach, requiring that we know the founders' pre-entrepreneurship occupation and industry codes. Appendix C.3 describes how we map LinkedIn job titles and firm names to occupation and industry codes using auxiliary data from Burning Glass and the Department of Labor's O-NET database. This match is imperfect, as it relies on unstructured data, so we account for uncertainty using multiple approaches that allow us to examine sensitivity.

We use the 2005 1% ACS sample from IPUMS, as 2005 is the first year that the ACS instituted a sample larger than 0.5% of the population Ruggles et al. (2021). From the 2005 1% ACS sample, we compute the probability that an individual drawn at random from each 3-digit Standard Occupation Classification (SOC) code - by - 2 digit NAICS code - by - age group {20-35, 36-51, 52+} cell is above a high-earner income threshold.[21] We then merge this probability into the LinkedIn data.[22] We then sum across founders to get the

[21]We use the 2005 ACS survey rather than a panel because the ACS top-codes earnings to protect confidentiality, and the earnings top-coding is different across years of the ACS. We believe it is conservative to rely primarily on across-industry and occupation variation in the pre-founding salary that comes from a single year so that changes in top coding procedures across years do not drive inference about founder selection and earnings. Moreover, salaries at the occupation-year level are extremely strongly correlated across years. Given the constraints and the stability of earnings over a decade, we felt that the best and most transparent approach is to focus on one cross section of the ACS rather than the panel.

[22]This exercise requires that we observe occupation. In some sensitivity analyses, if we cannot classify a

number of high-earners by industry and year. Our estimation dataset contains the number of high-earner founders in the LinkedIn sample in each of the 27 VentureSource industry groups between 2001–2016.

While we do not observe wealth in the ACS, our model predicts that high-earners' entry rates should increase when expected time to product falls, as more founders will meet the lower wealth thresholds that rationalize entry. To conceptualize this, we test whether those above different earnings levels (our preferred one is $225,000) are more likely to become VC-backed founders in cloud-affected industries after the cloud-computing shock lowers expected time to product.[23]

Note that while our stylized model of financial returns has predictions for the level of entrepreneurship in the pre-period, we have abstracted from many important components of entry decisions across sectors, including the possibility that healthcare, consumer products, industrial, and energy industries might offer different compensating differentials to high-ability founders (Stern 2004). Our test is therefore about changes in high-earning founders, which can be observed in Panel B of Table 6. In this panel, we present the estimated number of founders earning above $225,000 for the VentureSource universe of firms, which we compute by weighting the matched LinkedIn data to reflect the total amount of entry. Between the pre- and the post-periods, there are 23 more high-earning founders annually in cloud affected industries relative to an annual baseline pre-period average of 33 founders, an increase of nearly 70%. There were about 625 startups annually in cloud-affected industries that received VC-funding in the pre-period, meaning that the 23 founder increase among those with high-earnings in a prior job represents about 4 percent of the pre-period founders.

founder's prior occupation or it is unreported, we assign founders to low earning occupations. In other cases, we ignore founders with missing occupations and reweight the data we do utilize to match the VentureSource universe. For founders where we do not observe industry or age, we use the subset of cells that we observe. We have explored various approaches to handle missing data, none of which substantively change our conclusions because we utilize difference-in-differences regressions that remove the common effect of data handling choices across cloud and non-cloud industries. The stability of our estimates suggest data limitations, and our approaches to address them, do not systematically vary between cloud and non-cloud industries in the pre- and post-periods.

[23]This test assumes that there is no differential trend in the pre-entry salaries of those entering industries impacted by cloud computing versus not. Holding fixed the pre-period distribution of occupations and industries, we find that differential trends in the average and the median pre-entry compensation for individuals entering cloud vs. non-cloud impacted startups are small. If anything, the pre-entry earnings rose slightly more for those entering industries benefiting from cloud computing towards the end of our sample period, which should make our results even stronger.

While our model predicts a modest increase in high-earners in non-cloud affected industries, there is actually a decline of 1 founder per year (3% of the pre-period level).[24] The difference across types of industries suggests that founding did not become uniformly more attractive for high-earners, but instead that cloud affected industries became more attractive relative to non-cloud industries, where founder counts among high earners were roughly unchanged.

We turn now to a formal difference-in-differences test of these patterns, which allows us to present different methods of classifying high-earners, different high-earner thresholds, and different sample compositions to account for the possibility that serial entrepreneurship is driving our results.

4.3 Difference-in-Differences Regressions of the Number of High-Earning Founders

Our formal difference-in-differences regressions are estimates from models of the form

$$N_{High,j,t} = \beta \text{Post 2006 x Cloud Affected} + \alpha_j + \delta_t + \varepsilon_{j,t} \tag{4}$$

where the parameter β captures the change in the count of high-earning founders in cloud affected industries in the post period, while α_j and δ_t are industry and time fixed effects. We reweight all data to match the VentureSource distribution of startups for each of 27 VentureSource industry by founding-year cells. We cluster standard errors by VentureSource industry.

Table 7 presents the results for three different high-earner thresholds, while displaying two different methods for matching between the LinkedIn data and the ACS. Regardless of the threshold or matching approach, we find robust evidence that the number of high-earning founders increases in cloud-affected industries in the post-period. Columns 1-3 present results using matching Method 1, which assigns the lowest number of high-earning founders, while Columns 4-6 assign the highest number.[25] When the entry threshold is $225,000, which

[24] It is possible that the improved relative attractiveness of cloud affected industries draws in potential founders from non-cloud affected industries in the post-period.

[25] We match founders' job titles to both Burning Glass job titles and the Department of Labor's O-NET job title file to get SOC codes. In Method 1, founders with missing data or SOC codes that we cannot classify are assigned to a below-median earnings occupation. In addition, because the match between founder job-titles and SOC codes is noisy, we use the lowest earnings cell when matches are ambiguous. That is, if the best

is our preferred specification based on the theoretical analysis, we find that the number of high-earning founders increases by between 1.9 to 3.7 annually in each of the 8 cloud-affected industries in the post-period. These estimates represent increases of 60-74% relative to the pre-period number of high-earning founders (displayed below each column). We find similar qualitative results under high earner thresholds of \$150,000 and \$300,000. We do not examine thresholds above \$300,000 because the ACS data top-codes income to preserve confidentiality.

We note that the number of high-earning founders may be higher or lower than what is reported in the table, as our matching on observables approach does not account for positive or negative selection on unobservables among the actual founders who enter. What is important for our differences-in-differences identification is that these unobservables do not change differentially across cloud and non-cloud industries over time, allowing us to make relative comparisons. One possible violation of this identifying assumption is that our LinkedIn data quality or sampling frame may introduce differential changes for high-earners across sectors. Appendix Table C2 explores this possibility by looking at changes in an auxiliary measure of founder quality that we observe for the universe– the amount of capital invested in the first round of financing – to assess whether the LinkedIn sample's representation vis-à-vis the VentureSource universe changes across cloud and non-cloud industries in the pre- and post-periods. We find no evidence that changes in LinkedIn coverage drive our results.

Two other interpretation issues are relevant for this analysis. The first is verification of the parallel trends assumption. Figure 6 Panel A provides evidence on this assumption using a dynamic specification. We cannot reject zero differences between cloud and non-cloud industries with either classification method in the pre-period (p-values are 0.34 and 0.28). We then see clear breaks in cloud affected industries for firms founded after 2008, when cloud computing diffusion begins to ramp up. The post-period point estimates are jointly different from zero at the 1% level.

The second interpretation issue surrounds serial founders. Over 30% of founders in

match differs when using Burning Glass versus O-NET as the intermediate data step in matching to SOC codes, we assign the SOC code with the lowest average earnings among the best matches. There are 644 founders with SOC codes that we cannot classify, 711 with missing experience from their LinkedIn profile despite being old enough to have labor market experience, and 2,990 founders with ambiguous SOC codes. Method 2 ignores missing experience and assigns multiple matches to the SOC code with the higher average earnings.

LinkedIn have a founder or co-founder title immediately prior to their current venture. It is possible that faster time to product or failure in cloud affected industries simply means we are picking up more churn of serial entrepreneurs who exit failed ventures and start new ones. Figure 6 Panel B accounts for this possibility by dropping entrepreneurs with founder or co-founder titles prior to their venture in question. We continue to see strong evidence of an increase in high-earning founders in cloud affected industries relative to non-cloud industries in the post-period, with post-period point estimates that are jointly significant using Method 2 and pre-period estimates that are jointly insignificant for both specifications. We note that the implied number of additional founders is smaller (and estimates are less precise) when we examine only serial entrepreneurs, as we lose the potential for more than 30% of the total founders to be classified as high-earners. As our regressions are about counts, we would expect smaller point estimates.

Together, these analyses that use the career histories of those who became founders to study realized selection into VC-backed entrepreneurship provide strong validation of the model's predictions. They highlight how entry into entrepreneurship among the highest earning individuals in the workforce is sensitive to changes in the expected time to product. Because of this, industry-level variation in the expected time to product not only predicts differences in entry rates, but also has an impact on the composition of entrants across the pre-entry earnings distribution. Those in the highest percentiles of earnings in the workforce are disproportionately more likely to start ventures in industries where the expected time to product is faster. To the extent that earnings are correlated with ability, this is likely to impact the distribution of entrepreneurial talent across industries. Moreover, the fact individuals' backgrounds play an important role in shaping the startups they found, this is also likely to have implications for the supply of innovative ideas across industries.

5 Conclusion

In this paper, we combine several novel sources of data with a dynamic model of selection into entrepreneurship to study the length of time founder-CEOs hold non-diversifiable risk and the implications this has for constraints in the supply of entrepreneurial talent to VC-backed startups. We show that although founder-CEOs start with a low salary, the increases in CEOs' cash compensation after a startup begins product development are large enough that non-diversifiable risk held by founder-CEOs falls substantially after achieving this milestone. In turn, this implies that a startup's expected time to developing its first product is a key determinant of the attractiveness of VC-backed entrepreneurship.

Based on the differential reductions in expected time to product following the introduction of cloud computing services for software and internet startups, our model predicts that twice as many individuals in the highest 2% of the pre-entry earnings distribution would have an entrepreneurial entry constraint relaxed in industries that benefited from cloud-computing relative to those not impacted by cloud-computing. Empirical analysis of founder backgrounds and estimates of pre-founding earnings provides strong support for the theoretical predictions. In doing so, we provide some of the first evidence for the particular sectors where the supply of entrepreneurial talent, and hence VC-backed innovation, may be most likely to be constrained.

Our results highlight an under-appreciated role played by venture capital investors– as that of intermediate liquidity providers – which they might be uniquely positioned to do as hands-on investors. This role of venture capitalists as liquidity providers may also help explain the sectors where VCs are more actively involved in financing innovation. The faster resolution of uncertainty for firms in certain sectors implies a lower need for VCs to provide intermediate liquidity to founders in order to induce them to become entrepreneurs. Sectors where uncertainty is resolved more slowly are those where the burden of non-diversifiable risk is likely to be most extreme. The degree to which these individuals' ideas are not commercialized (or commercialized inside incumbent firms) as well as the aggregate impact of this selection remains an interesting area of further work.

References

Aghion, Philippe and Patrick Bolton. 1992. "An incomplete contracts approach to financial contracting." *The Review of Economic Studies* 59 (3):473–494.

Anton, James and Dennis A Yao. 1995. "Start-ups, Spin-offs, and Internal Projects." *Journal of Law, Economics, and Organization* 11 (2):362–78.

Babina, Tania, Paige Ouimet, and Rebecca Zarutskie. 2017. "Going entrepreneurial? IPOs and new firm creation." *Working Paper* .

Baker, Malcolm P and Paul A Gompers. 1999. "Executive ownership and control in newly public firms: The role of venture capitalists." *Working paper* .

Bengtsson, Ola and John RM Hand. 2011. "CEO compensation in venture-backed firms." *Journal of Business Venturing* 26 (4):391–411.

Bhidé, Amar V. 2003. *The Origin and Evolution of New Businesses.* Oxford University Press.

Botelho, Tristan L and Melody Chang. 2022. "The Evaluation of Founder Failure and Success by Hiring Firms: A Field Experiment." *Organization Science* .

Díaz-Giménez, Javier, Andy Glover, and José-Víctor Ríos-Rull. 2011. "Facts on the distributions of earnings, income, and wealth in the United States: 2007 update." *Federal Reserve Bank of Minneapolis Quarterly Review* 34 (1):2–31.

Dillon, Eleanor and Christopher Stanton. 2018. "Self-employment Dynamics and the Returns to Entrepreneurship." *Working Paper* .

Edmans, Alex, Xavier Gabaix, and Dirk Jenter. 2017. "Executive compensation: A survey of theory and evidence." In *The Handbook of the Economics of Corporate Governance*, vol. 1. Elsevier, 383–539.

Edmans, Alex, Xavier Gabaix, Tomasz Sadzik, and Yuliy Sannikov. 2012. "Dynamic CEO compensation." *The Journal of Finance* 67 (5):1603–1647.

Ewens, Michael and Joan Farre-Mensa. 2020. "The deregulation of the private equity markets and the decline in IPOs." *The Review of Financial Studies* 33 (12):5463–5509.

Ewens, Michael and Matt Marx. 2017. "Founder replacement and startup performance." *The Review of Financial Studies* 31 (4):1532–1565.

Ewens, Michael, Ramana Nanda, and Matthew Rhodes-Kropf. 2018. "Cost of experimentation and the evolution of venture capital." *Journal of Financial Economics* 128 (3):422–442.

Gompers, Paul, Josh Lerner, and David Scharfstein. 2005. "Entrepreneurial spawning: Public corporations and the genesis of new ventures, 1986 to 1999." *The Journal of Finance* 60 (2):577–614.

Gompers, Paul Alan and Joshua Lerner. 2004. *The venture capital cycle*. MIT press.

Gornall, Will and Ilya A Strebulaev. 2020. "Squaring venture capital valuations with reality." *Journal of Financial Economics* 135 (1):120–143.

Hall, Robert E and Susan E Woodward. 2010. "The Burden of the Nondiversifiable Risk of Entrepreneurship." *American Economic Review* 100 (3):1163–1194.

Hellmann, Thomas. 2007. "When do Employees Become Entrepreneurs?" *Management Science* 53 (6):919–933.

Hellmann, Thomas F, Paul H Schure, Chloe Tergiman, and Dan H Vo. 2019. "Ownership dynamics within founder teams: The role of external financing." *Strategic Entrepreneurship Journal* 13 (3):256–287.

Kerr, William R, Ramana Nanda, and Matthew Rhodes-Kropf. 2014. "Entrepreneurship as experimentation." *The Journal of Economic Perspectives* 28 (3):25–48.

Kim, J Daniel. 2018. "Is there a startup wage premium? Evidence from MIT graduates." *Research Policy* 47 (3):637–649.

Koning, Rembrand, Sampsa Samila, and John-Paul Ferguson. 2021. "Who Do We Invent for? Patents by Women Focus More on Women's Health, but Few Women Get to Invent." *Science* 372 (6548):1345–1348.

Kortum, Samuel and Josh Lerner. 2000. "Assessing the contribution of venture capital to innovation." *RAND Journal of Economics* 31 (4):674–692.

Lazear, Edward P. 2004. "Balanced skills and entrepreneurship." *American Economic Review* 94 (2):208–211.

Lee, David S. 2009. "Training, wages, and sample selection: Estimating sharp bounds on treatment effects." *The Review of Economic Studies* 76 (3):1071–1102.

Lerner, Josh and Ramana Nanda. 2020. "Venture Capital's Role in Financing Innovation: What We Know and How Much We Still Need to Learn." *Journal of Economic Perspectives* 34 (3):237–261.

Manso, Gustavo. 2016. "Experimentation and the Returns to Entrepreneurship." *Review of Financial Studies* 29:2319–2340.

Moretti, Enrico. 2010. "Local labor markets." *NBER Working Paper* .

Rajan, Raghuram G. 2012. "Presidential address: The corporation in finance." *The Journal of Finance* 67 (4):1173–1217.

Ruggles, Steven, Sarah Flood, Ronald Goeken, Josiah Grover, Erin Meyer, Jose Pacas, and Matthew Sobek. 2021. "Ipums usa: Version 10.0 [dataset]. Minneapolis, MN: IPUMS, 2020."

Saez, Emmanuel and Gabriel Zucman. 2016. "Wealth inequality in the United States since 1913: Evidence from capitalized income tax data." *The Quarterly Journal of Economics* 131 (2):519–578.

Silva, JMC Santos and Silvana Tenreyro. 2006. "The log of gravity." *The Review of Economics and Statistics* 88 (4):641–658.

Stern, Scott. 2004. "Do scientists pay to be scientists?" *Management science* 50 (6):835–853.

Wasserman, Noam. 2003. "Founder-CEO Succession and the Paradox of Entrepreneurial Success." *Organization Science* 14 (2):149–172.

Figure 1: Founder-CEO Cash Compensation by Firm Age and Capital Raised

This figure displays founder-CEO cash compensation by firm age and capital raised. The left panels include all firms and the right panels restrict the sample to firms that are still in the product definition or ideation phase. Firm age for pre-product firms ends at 4 because there are no older pre-product firms in the AHR data. There are also no pre-product firms with over $100 million in venture capital raised.

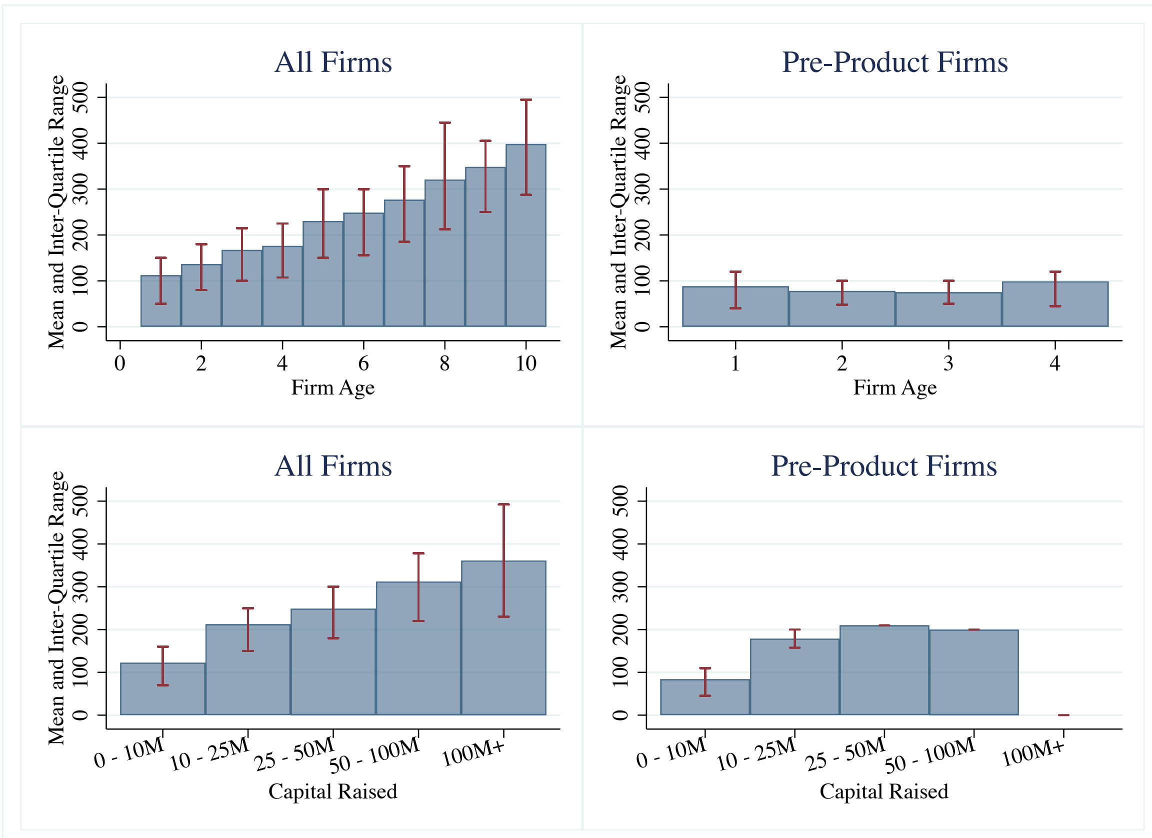


Figure 2: Predicted Densities of Founder-CEO Cash Compensation by Product Status and Revenue.

This figure displays density plots of predicted cash compensation in levels, taken from the Poisson regressions in Equation (1). Density plots of predicted cash compensation are by product and revenue status. “Pre-product” are the set of firms without a product and “Post-Product, pre-revenue” are those firms with a product but no reported revenue. The categories “0 to 25 million” and “25 to 100 million” are firms with those ranges of reported revenue. To improve readability, firms with greater than $100 million in revenue are not included in the plots despite being in the regression.

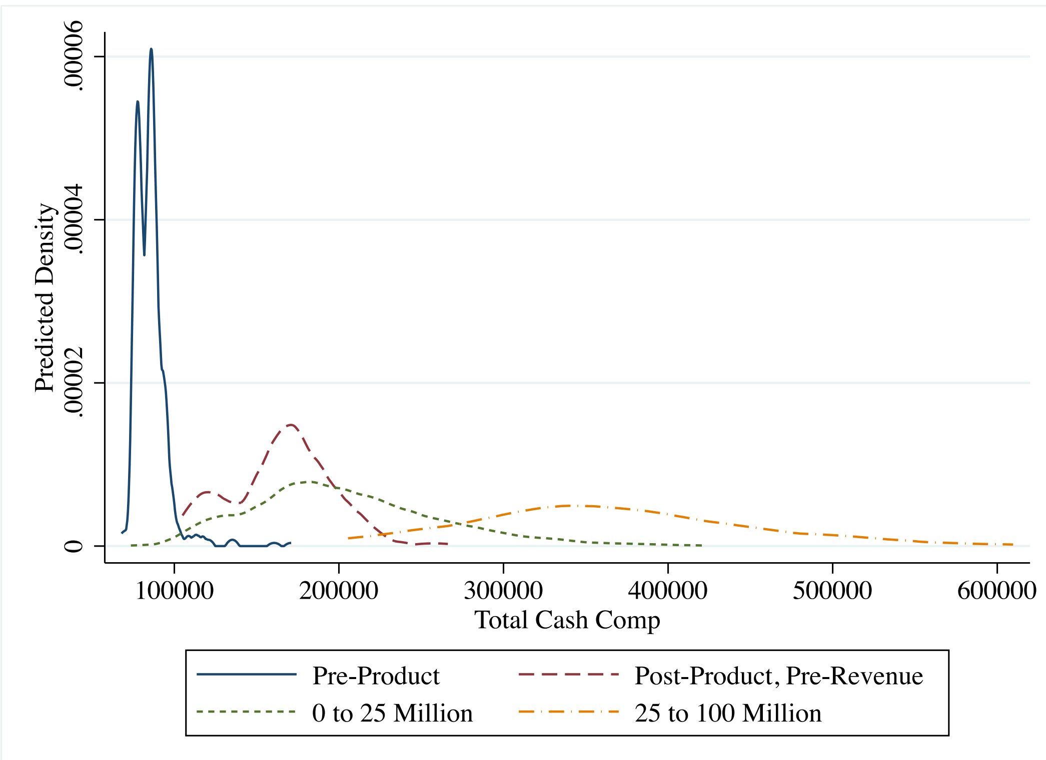
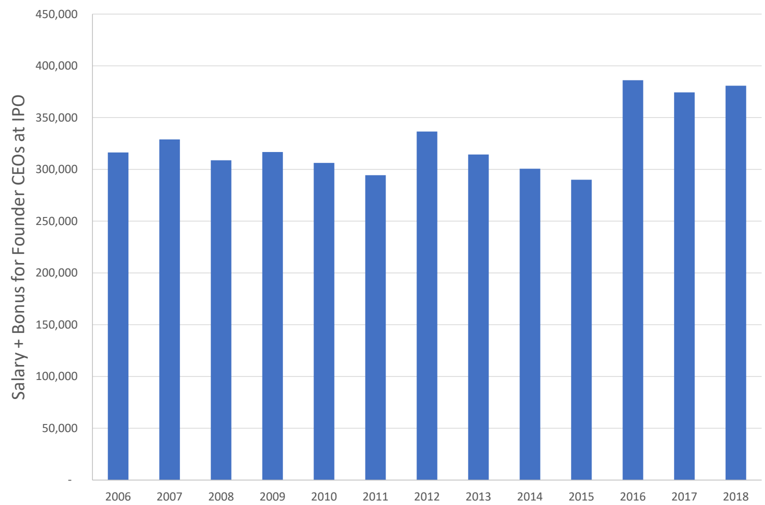


Figure 3: Validating AHR's Post-Product CEO-Compensation Statistics using Public Company Data

Panel (a) reports the median salary plus bonus for founders of newly-public, formerly VC-backed startups, in constant 2015 dollars. The compensation data is extracted from S-1 filings when available and otherwise collected from DEF 14-A filings. Data construction details are founder in Appendix A.3. For panel (b), data on private firms come from the AHR survey. Compensation data of public firm CEOs is taken from Execucomp for U.S.-based public firms found in Compustat. We use the scraped 14-A filings from (a) to fill in any firm-year gaps in Execucomp (it covers larger public firms). For public firms, we drop financials and utilities and exclude CEOs with under \$5 in salary or under \$5 in total cash compensation. The sample of public firms in the compensation data over-weights large firms relative to the Compustat universe of publicly traded firms, so we re-weight the compensation data to reflect the Compustat universe.

(a) Founder-CEO Salary+Bonus at IPO



(b) Relationship between Firm Revenue and CEO-Compensation: Public Firms vs. AHR data

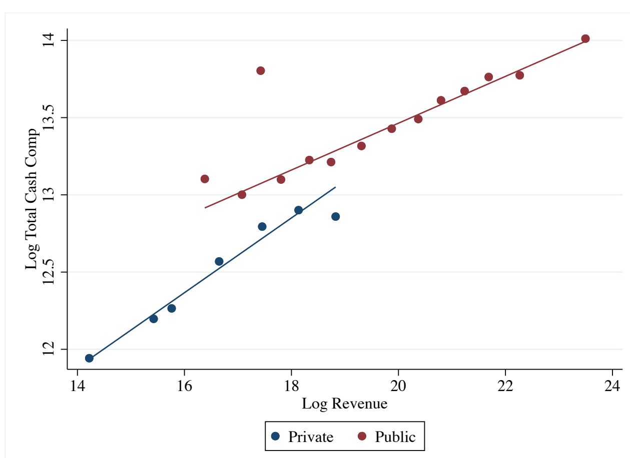


Figure 4: Attractiveness of Founding based on Compensation Contract and Time to Product

This figure compares the certainty equivalent of entrepreneurship under a fixed salary of $150,000 over the life of a venture (top line) to a contract where compensation changes with product status and firm age (middle line). The product- and age-based compensation contract comes from moments of the Advanced HR data. The bottom line uses the product and age-based compensation contract but eliminates the pre-product phase, showing the region with incentive-compatible entry if the founder were able to enter the post-product phase immediately. The coefficient of relative risk aversion is assumed to be 2. The area above each line is the region where the certainty equivalent is positive and founding is incentive-compatible. Percentiles of the income distribution are displayed below the x-axis, which we compute from the Statistics of Income (SOI) data stored at the NBER. We use data items 85 and 86 from the SOI data, which contain W2 earnings for individual filers and married joint filers. Individuals without W2 earnings are not included in the percentile estimates. Wealth percentiles on the right y-axis are interpolated from data provided in Saez and Zucman (2016).

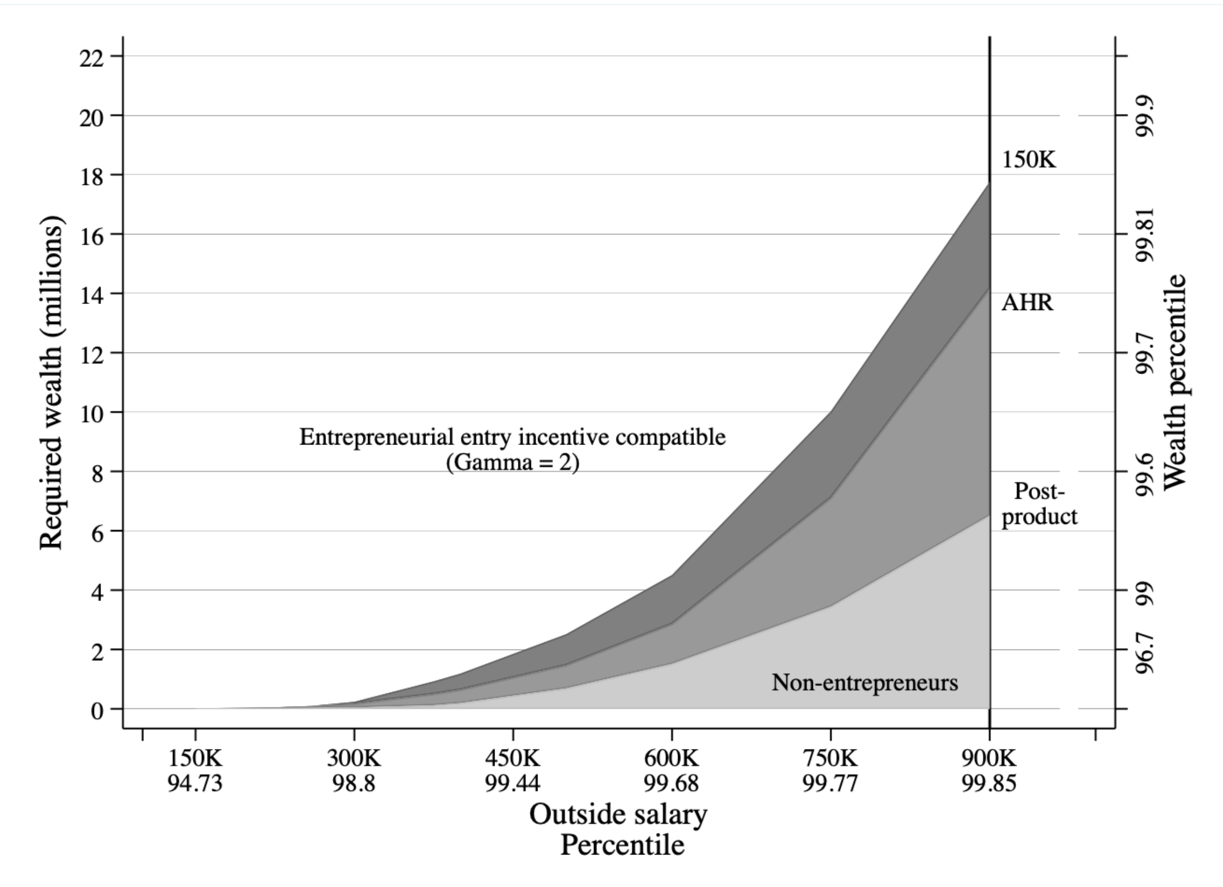
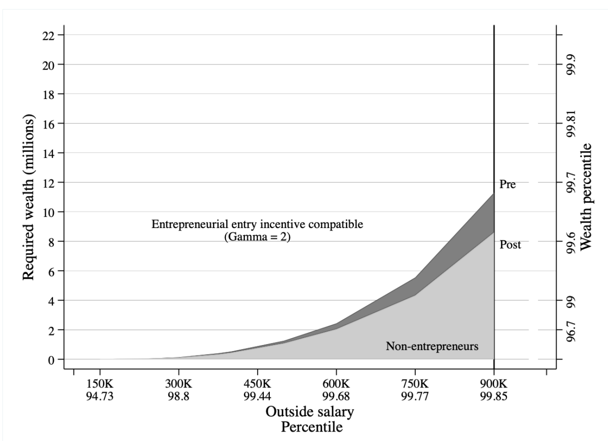
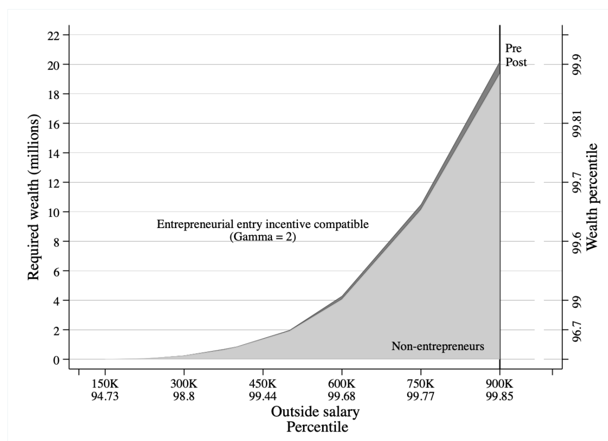


Figure 5: Cross-Industry Heterogeneity in Time to Product and Attractiveness of VC-Backed Entrepreneurship

Notes: This figure compares salary-wealth thresholds such that a potential entrepreneur may find founding attractive. Each panel compares these thresholds under the pre- and post-period product transition probabilities and pre-product exit probabilities given in Appendix Table B1. Panel (a) displays the impact of faster time to product between the pre- and post-periods in cloud-affected industries. In cloud-affected industries, 15% more firms achieve product status resolution by year 3 in the post-period compared to the pre-period. Panel (b) displays the results for non-cloud affected industries, where there were minimal changes in time to product. The area above each line is the region where the certainty equivalent is positive and founding is incentive-compatible. In each sub-figure, the top line corresponds to the salary-wealth threshold for entrepreneurial entry to be incentive compatible using the Pre-2006 product transition and process for the industry. The bottom figure uses the Post-2006 product transition process for the industry. As discussed in the text, we also adjust founders' equity holdings in the post-period in Panel (a) to reflect that faster time to product is associated with less dilution, increasing founders' equity stakes by about 8 percent, or 1.75 percentage points. Comparable figures that hold fixed the distribution of exit values across industries and equity values across time are given in Appendix Figure B2. Percentiles of the income distribution are displayed below the x-axis, which we compute from the Statistics of Income (SOI) data stored at the NBER. We use data items 85 and 86 from the SOI data, which contain W2 earnings for individual filers and married joint filers. Individuals without W2 earnings are not included in the percentile estimates. Wealth percentiles on the right y-axis are interpolated from data provided in Saez and Zucman (2016).



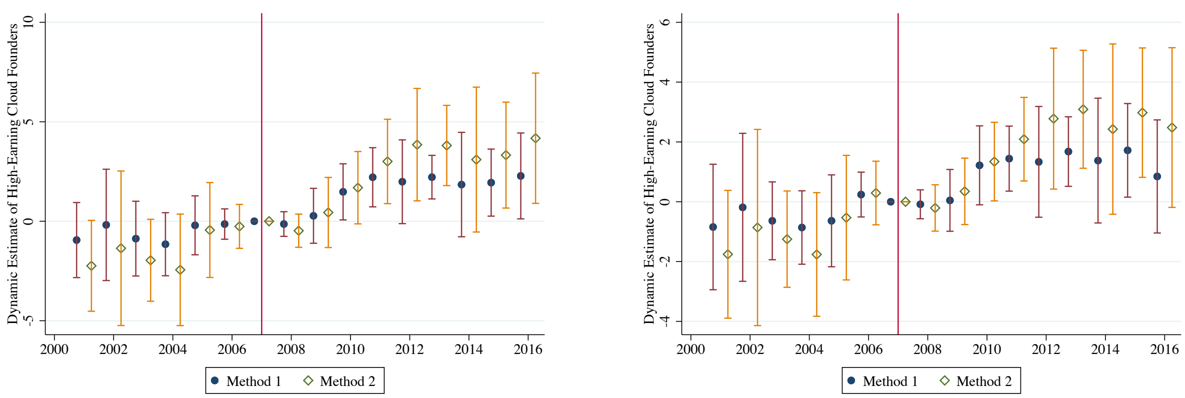
(a) Cloud Product Transitions, Cloud Sample



(b) Non-Cloud Transitions, Non-Cloud Sample

Figure 6: Dynamic Difference-in-Differences Estimates of High-Earning Founder Entry Into Cloud Affected Industries

Notes: This figure reports dynamic difference-in-differences estimates of the differential number of high-earning founders entering cloud affected industries compared to non-cloud affected industries. The dependent variable is the sum of all observed founders' probabilities of earning over $225,000 in their pre-founding job. This probability is based on a match of LinkedIn data on founder backgrounds to the share of the population that earns more than $225,000 in the founders' prior 3-digit SOC code, 2-digit NAICS, and coarsened age cell using the 2005 American Community Survey data for those with a bachelors or higher degree. When age is unobserved, we take the 3-digit SOC and 2-digit NAICS cell. We then sum the probability that an entering founder is a high-earner at the industry-year level for 27 VentureSource industries. We weight industries-years in the LinkedIn data to match the VentureSource universe. Because the mapping from LinkedIn job titles to occupation data is noisy, we report estimates using two methods. In Method 1, when a match is ambiguous, we assign an occupation or industry that has the lowest average earnings. When background or experience is missing, we assign a below-median earnings cell. In Method 2 we ignore missing experience and assign ambiguous matches to the SOC code with the highest average earnings. Confidence intervals are based on standard errors clustered by industry. Panel A presents estimates for the full sample while Panel B drops serial entrepreneurs who have "founder" anywhere in their pre-entry job title. Wald tests never reject the pre-period estimates are different from zero, while always rejecting that the post-period estimates are jointly zero.



(a) Full Sample

(b) Dropping Serial Entrepreneurs

Table 1: Descriptive Statistics for Founder-CEOs

Notes: Data come from a survey of startups who are encouraged to participate by their investors. The sample includes cross-sections for U.S. based founder CEOs in technology (consumer, enterprise, hardware, other technology) firms for 2015 ($N = 687$), and 2017 ($N = 1289$). Changing numbers of observations across years arise because 1) there are more investments, especially in seed and early-stage rounds, and 2) new funds are added to the survey (although top tier VC firms are included in every survey wave). Cell means are reported on pooled survey years. The survey protects anonymity by coarsening revenue, number of employees (headcount), and total capital raised. Panel B reports data by coarsened revenue category. Among pre-revenue firms we distinguish those that are "Pre-Product" and those that are "Post-Product." Pre-Product firms are those pre-revenue firms that report "Early / Product Definition" as their development stage. Post-Product firms report "Product Development," "Product in Beta," "Shipping Product" or "Profitable" as their product development stage. All firms with strictly positive revenue are "Post-Product" firms. Average cumulative venture capital raised and average headcount are averages of the mid-points coarsened categories in the survey. For averages involving highest coarsened category, we use the conditional mean for the category calculated from uncensored data in VentureSource. The survey reports the exact value of each executive's salary, target bonus, and fully diluted equity in the firm. Total target cash compensation is the sum of salary and target bonus. CEO's equity ownership is calculated using fully diluted equity in the firm on an as-converted basis, including any option pools and convertible preferred stock. Fully diluted equity is reported for the CEO, it is not the total employee or co-founders' equity in the firm.

	Share of firms	Firm age (years)	Cumulative Venture Capital Raised ($ M)	Average Headcount	Mean of CEO's Total Target Cash Compensation	Std. Dev of Total Target Cash Compensation	CEO's Equity Ownership
Panel A: By round of financing							
Seed	32%	2.2	3.3	10	97,969	68,434	38%
Series A	27%	3.6	13.7	39	179,624	94,067	22%
Series B	20%	4.9	38.9	87	238,160	111,302	16%
Series C	11%	6.2	82.7	148	299,339	148,891	13%
Series D	6%	7.7	123.7	191	353,550	121,749	11%
Series E and beyond	4%	9.1	205.9	241	403,433	159,236	8%
Panel B: By revenue							
Pre-Product, Pre-Revenue	22%	1.8	3.8	7	85,020	62,619	40%
Post-Product, Pre-Revenue	9%	2.3	19.5	25	165,330	78,982	24%
$0-$10M	42%	4.1	23.3	46	186,898	93,881	21%
$10M-$25M	12%	6.2	59.3	117	265,516	115,012	15%
$25M-$50M	6%	7.1	93.9	192	333,789	135,798	14%
$50M-$100M	5%	7.9	130.1	219	402,382	152,148	13%
$100M+	3%	8.0	164.1	274	382,422	232,817	15%

Table 2: Poisson Regressions of Cash Compensation on Milestones for VC Backed Founder-CEOs

Notes: This table reports Poisson regressions of cash compensation (salary plus bonus) on firm characteristics and milestones from the AHR survey. The sample is restricted to founder-CEOs. Survey data are coarsened to protect firm anonymity, so we use indicators for different milestone categories. The regression in Column 6 is re-weighted to reflect the VentureSource distribution of firm age and capital raised within each survey year. Standard errors in parentheses are clustered by founding year-by-region-by-industry cell. Post Series D and Post Series E (omitted due to space) are similar to Post Series C. The Post Product Definition indicator is a dummy that the firm has moved past early stage product definition into: product development, a beta product, shipping product, or profitable sales.

	(1)	(2)	(3)	(4)	(5)	(6)
Post Product			0.616***	0.496***	0.365***	0.363***
			(0.044)	(0.043)	(0.051)	(0.062)
Revenue (Baseline is Pre-Revenue)						
$0M-$10M	0.543***	0.401***	-0.054	-0.039	-0.074	-0.014
	(0.039)	(0.038)	(0.046)	(0.045)	(0.045)	(0.048)
$10M-$25M	0.894***	0.609***	0.152*	0.049	0.001	0.095
	(0.046)	(0.051)	(0.054)	(0.052)	(0.051)	(0.065)
$25M-$50M	1.120***	0.783***	0.329***	0.192**	0.152*	0.221**
	(0.051)	(0.056)	(0.063)	(0.061)	(0.059)	(0.060)
$50M-$100M	1.309***	0.945***	0.490***	0.327***	0.291***	0.408***
	(0.066)	(0.069)	(0.077)	(0.078)	(0.077)	(0.102)
$100M+	1.261***	0.902***	0.445***	0.267**	0.273**	0.188
	(0.085)	(0.082)	(0.083)	(0.086)	(0.089)	(0.114)
VC Funding Round (Seed is Baseline)						
Post Series A					0.263***	0.230***
					(0.043)	(0.049)
Post Series B					0.377***	0.306***
					(0.064)	(0.064)
Post Series C					0.455***	0.359***
					(0.073)	(0.078)
Firm Age Dummies		Yes	Yes	Yes	Yes	Yes
Cumulative VC Raised Dummies				Yes	Yes	Yes
Region and Industry Dummies				Yes	Yes	Yes
Unreported Venture Round Dummies					Yes	Yes
Re-weighted						Yes
Pseudo R-Squared	0.390	0.441	0.479	0.507	0.527	0.419
Observations	1,976	1,976	1,976	1,976	1,976	1,975

Table 3: Summary Statistics on Founder vs. Non-Founder CEOs of VC Backed Firms

Notes: Descriptive Statistics are drawn from the proprietary AHR survey data and pool years 2015 and 2017. Non-founders make up 20.5% of the overall sample.

		Founder CEOs				Non-Founder CEOs			
Firm Revenue	Share of CEOs that are founders or co-founders	CEO's Total Target Cash Compensation	CEO's Base Salary	Base Salary Share of Cash Compensation	CEO's Equity Ownership	CEO's Total Target Cash Compensation	CEO's Base Salary	Base Salary Share of Cash Compensation	CEO's Equity Ownership
Pre-Product, Pre-Revenue	98%	85,020	80,264	94%	40%	86,886	86,886	100%	13%
Post-Product, Pre-Revenue	91%	165,330	158,662	96%	24%	309,403	252,339	82%	7%
\$0-\$10M	83%	186,898	173,982	93%	21%	347,876	277,574	80%	6%
\$10M-\$25M	67%	265,516	232,021	87%	15%	406,546	306,610	75%	5%
\$25M-\$50M	53%	333,789	277,579	83%	14%	456,903	329,827	72%	5%
\$50M-\$100M	61%	402,382	308,999	77%	13%	529,358	358,834	68%	4%
\$100M+	64%	382,422	316,696	83%	15%	569,796	380,341	67%	5%

Table 4: Analysis of Founder-CEO Turnover

Notes: This table reports regressions of a dummy for a non-founder CEO on the post product dummy and other measures of the firm lifecycle. Columns 1-3 use AHR data and Columns 4-6 use data from VentureSource for industries close to those covered by AHR. The AHR sample is conditioned on firms that are alive at the time of the survey. The founder-CEO replacement information in Columns 4-6 builds on the data from Ewens and Marx (2017), where we infer the timing of CEO replacement from quarterly downloads available since 2011Q4, which we supplement with the LinkedIn data on founders (see Appendix C.3). In Columns 4 and 5 we consider the sub-sample of years where startups are alive in 2015 and 2017. "Post-product Definition" for this sample is equal to one if VentureSource reports that the company is "Generating Revenue," is "Profitable," or has a "Product in Beta Test." Pre-product firms are those that VentureSource reports as "Startup" or in "Product Development." These statuses are observed at every new financing event. Column 6 takes advantage of the full panel of startups first financed from 2000 to 2014 (to provide time for replacements to be observed through the end of sample in 2020Q1). Here a unit of observation is a startup-year, where startups are tracked from first VC financing to either exit, founder-CEO replacement or the end of sample (whichever comes first). This specification can be thought of as analogous to estimating the turnover hazard. All specifications have year fixed effects. Standard errors are clustered by founding year-by-region-by-industry cell in Columns 1-3 and by firm in Column 6. Results for the VentureSource sample are similar if we use all industries.

	AHR Sample			Venture Source		
	(1)	(2)	(3)	(4)	(5)	(6)
	All AHR Firms			Startups Alive in 2015 and 2017 backed by AHR VCs		Hazard Model for all startups
Post Product Definition	0.051***	0.052***	0.044*	0.040***	0.040***	0.019***
	(0.014)	(0.014)	(0.021)	(0.007)	(0.008)	(0.001)
Firm Age	Yes	Yes	Yes	Yes	Yes	Yes
Industry		Yes	Yes		Yes	Yes
Region		Yes	Yes			
Revenue			Yes			
Mean DV for Pre-Product Firms	0.016	0.016	0.016	0.012	0.012	
Mean of Pre-Product Entry Hazard						0.005
R-Squared	0.242	0.252	0.256	0.142	0.150	0.017
Observations	2,485	2,485	2,485	4,247	4,245	167,621

Table 5: Minimum Wealth Required for Founding to Have a Positive Expected Value Under Various Scenarios

This table displays wealth required (in millions of dollars) for founding to have positive expected utility. The first column displays the outside salary for a potential founder considering VC backed entrepreneurship. Each cell then displays the required minimum wealth to make founding a positive expected value choice for an agent with a coefficient of relative risk aversion of 2. The bottom row in each panel provides reductions in the number of missing founders under the assumption that income and wealth are bivariate lognormal with a correlation coefficient of 0.48. Columns 1 and 2 fit models using product timing and exit outcomes data for firms founded prior to 2006. The data are sparse for firms founded prior to 1990. Columns 3 and 4 use the pre-period exits data to calculate returns to founders under product transition processes estimated on firms founded between 2007 and 2014. Column 4 adds 1 year to exit for firms that IPO or have acquisitions that provide a return to common shareholders. In Columns 2-4, compensation moments come from the AHR data with an autocorrelation structure as described in the text. We estimate founder equity holdings using averages in AHR by firm-age for the Post-Period. We adjust equity holdings for the pre-period using the ratio of Compensation Pro equity holdings to AHR equity holdings by industry. In Panel A Columns 3 and 4, we assume that founder equity is responsive to speed to product and increases by 8% in the post-period (averaging 1.75 percentage points, to match the AHR post-period data), as Ewens, Nanda, and Rhodes-Kropf (2018) show that cloud-affected industries in the post-period featured smaller initial investments and larger valuation stepups upon success. Without the change in equity in the Post-2006 Cohorts in Panel A, the change in the number of missing founders in Column 3 is 290,669.

Panel A: Cloud Affected Industries

Outside Salary	Pre-2006 Cohorts (34.9% Pre-Product after 3 Years)		Post-2006 Cohorts (29.7% Pre-Product after 3 Years)	
	(1)	(2)	(3)	(4)
	Flat $150,000	AHR Moments	AHR Moments	AHR Moments, but add 1 year before successful exit
$150,000	0.00	0.02	0.02	0.02
$225,000	0.04	0.04	0.04	0.04
$400,000	1.08	0.53	0.45	0.46
$500,000	2.31	1.25	1.08	1.11
$600,000	3.91	2.43	2.04	2.17
$750,000	8.44	5.55	4.33	4.47
$900,000	15.37	11.29	8.60	8.81
Reduction in Missing High-Earning Founders Relative to Column (1)		251,602	312,587	300,945

Panel B: Non-Cloud Industries

Outside Salary	Pre-2006 Cohorts (56.1% Pre-Product after 3 Years)		Post-2006 Cohorts (55.6% Pre-Product after 3 Years)	
	(1)	(2)	(3)	(4)
	Flat $150,000	AHR Moments	AHR Moments	AHR Moments, but add 1 year before successful exit
$150,000	0.00	0.02	0.02	0.02
$225,000	0.05	0.07	0.07	0.07
$400,000	1.38	0.86	0.83	0.83
$500,000	3.03	1.97	1.91	1.93
$600,000	6.26	4.29	4.06	3.94
$750,000	13.39	10.50	10.13	9.76
$900,000	23.78	20.21	19.37	19.22
Reduction in Missing High-Earning Founders Relative to Column (1)		170,247	194,047	190,996

Table 6: Backgrounds of Founders of VC-backed Startups

This table summarizes biographical information for the founders and early CEOs of startups where we could find their Linkedin profile. The table presents the summary statistics for the full sample of founders at the time of founding along with splits by industry (cloud vs. non-cloud). Age at founding is the age inferred from an individual's undergraduate education start date (assuming they are 18 at the time of entry). # jobs pre-founding" reports the unique number of job titles prior to the founding event. # years pre-founding experience" is the number of years with work experience prior to founding. "Founder of Startup before?" is equal to 1 if the founder has "founder" in any pre-founding titles and "CEO pre-founding" is 1 if there is a CEO title in the pre-founding experience. The next set of variables report the highest level of education achieved by the time of founding.The "Pre-founding: " variables report whether the founder had any of the other categories. Panel B reports the estimated number of founders who enter annually from the top 2% of the earnings distribution. These estimates come from the probability that a founder who is randomly drawn from the empirical distribution of those in the same 3-digit occupation code, 2-digit NAICS code, and age cell earns above $225,000 in the 2005 ACS. When occupation codes are ambiguous, we average the probability the founder is a high earner across possible matches. When industry or age are missing, we use cell averages for the non-missing data. These estimates are weighted to match the VentureSource universe. Matching process details are described in the text and in Appendix C. The numbers in Panel B are averages of the Method 1 and Method 2 approaches in Table 7.

Panel A: Founder-CEOs with LinkedIn Backgrounds

	Full Sample	Entering industries impacted by cloud computing	Entering industries not impacted by cloud computing
Number of Founders	9,574	7,051	2,523
Age at founding	34.5	33.5	37.2
Number of years in Labor force	11.7	11.3	12.8
Number of jobs pre-founding	4.4	4.5	4.2
Share founded a startup before	36%	38%	30%
Bachelors Degree or higher	73%	73%	74%
MBA	24%	23%	26%
Non-MBA Masters	11%	11%	13%
MD and/or JD	9%	7%	13%
PhD	13%	8%	25%

Panel B: ACS Data using LinkedIn match

	Full Sample	Entering industries impacted by cloud computing	Entering industries not impacted by cloud computing
Estimated number of founder-CEOs entering each year from highest 2% of pre-entry earnings in pre-period	66	33	33
Estimated number of founder-CEOs entering each year from highest 2% of pre-entry earnings in post-period	88	56	32
Estimated increase in highest earning entrants per year	22	23	-1

Table 7: Difference-in-Differences Regressions of High-Earning Founder Entry

Notes: This table presents difference-in-differences estimates of the number of high-earning founders in cloud affected industries in the post-period (after 2006). The sample is based on counts of LinkedIn founders matched to VentureSource data in 27 VentureSource industry groups between 2001–2016. All data are re-weighted to match the VentureSource distribution of startups across industry and year. Counts of high-earning founders come from a probabilistic match based on the share of bachelors or higher degree holders in the the 2005 American Community Survey in an Occupation-Industry-Age cell that earns more than the target income threshold given in the column heading. For founders where we do not observe age, we use the Occupation-Industry cell average. We use 3-digit SOC codes, 2-digit NAICS codes, and age groups 20-35, 36-51, and 52+ to form the ACS to LinkedIn data match. Method 1, in Columns 1-3, assigns founders with missing experience or occupation data to a below-median earnings occupation. In addition, because the match between founder job-titles and SOC codes is noisy, when there are multiple matches we assign the SOC code with the lowest average earnings. Method 2, in Columns 4-6, ignores missing experience and assigns multiple matches to the SOC code with the highest average earnings. Below each column we report the industry-year average number of high-earning founders in the pre-period for 8 cloud-affected industries and 19 non-cloud affected industries. We also report the average number of new firms in the pre-period, by industry. Standard errors in parentheses are clustered by 27 VentureSource industries.

	Method 1			Method 2		
	(1)	(2)	(3)	(4)	(5)	(6)
Threshold	$150,000	$225,000	$300,000	$150,000	$225,000	$300,000
Post 2006 x Cloud Affected	4.413***	1.992***	1.477***	7.659***	3.741***	2.751***
	(1.146)	(0.474)	(0.351)	(1.989)	(0.907)	(0.666)
Industry FE	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Pre-Period Mean DV, Cloud	6.26	3.14	2.37	9.92	5.09	3.86
Pre-Period Mean DV, Non-Cloud	2.45	1.26	0.94	4.13	2.22	1.66
Avg New Firms in Pre-Period, Cloud	78.11	78.11	78.11	78.11	78.11	78.11
Avg New Firms in Pre-Period, Non-Clou	33.05	33.05	33.05	33.05	33.05	33.05
R-Squared	0.886	0.878	0.876	0.900	0.895	0.895
Observations	432	432	432	432	432	432

Internet Appendix for

Founder-CEO Compensation and Selection into Venture Capital-Backed Entrepreneurship

Appendix A Compensation Data

A.1 Data Construction and Validation

To assess possible response bias among the portfolio companies from the VC firms that participate in the AHR survey, we first restrict the universe of startups to those that have received any funding from one of these VCs. Table A1 displays this comparison between the AHR data and the eligible VentureSource startups.[26] While the response rate is high across survey waves, the number of firms rises in the 2017 wave because there are more startups receiving funding over time, and a larger number of VC investors participated in the 2017 survey.

Table A2 also examines whether the types of firms backed by VCs who participate in the survey are similar to firms backed by VCs who do not participate. It is evident that the AHR data is tilted somewhat towards firms that have raised more capital than those in the VentureSource universe. This difference largely come from the set of VC funds in the survey versus the universe of private investors. Participating VC funds in the 2017 survey had a median of $1.4 billion in assets under management, whereas non-participating VC funds had a median of $85 million under management. As mentioned in the main text, the funds in the survey managed 42% of total industry assets and deployed nearly 49% of the dollar-weighted investments in the VentureSource data.

A.2 Bounds on Sample Selection in the AHR Survey

To assess whether surviving, post-product firms observed at later ages are those that had high pay as pre-product firms, we trim the lower tail of the pre-product compensation distribution using the procedure in Lee (2009). This is akin to assessing how our results change if we assume that it is only the firms with the highest pre-product pay that survive to reach subsequent milestones. The ingredients for this calculation are relatively simple: conditional on being three years old and included in the AHR survey, 82% of firms have hit post-product status and 18% have not. In longitudinal data from VentureSource, a conservative estimate for our time period and sample

[26] We focus on firms born within 10 years of a survey year. To avoid so called 'zombie' firms, we drop firms that have not achieved an exit and have not raised financing within four years of their last funding event.

is that 20% of firms fail or exit by year three, meaning they are missing from the pay data by construction. Suppose 100 firms are born. By year three, 20% will not be present due to having exited or failed. Of the remaining 80, 18% (14.4 firms) remain pre-product, suggesting we should trim the lower $0.2 + .18 * .8 = 34.4$th percentile of the earnings distribution for pre-product firms under three years old. Let τ denote the trimming percentile of the pre-product earnings distribution for firms under three years of age. We then compute

$$\mathbb{E}(\log(Comp)|\text{Post-Product,Age=3}) - \mathbb{E}(\log(Comp)|\log(Comp) > \tau, \text{Pre-Product,Age} < 3).$$

as the bounds on sample selection.

A.3 Data used to Validate Robustness of AHR

One may ask whether our results using AHR and the sample period are representative of a longer time series, especially one where VC funding was less "frothy." To assess whether we are simply capturing an ephemeral moment in VC funding history, we bring to bear hand-collected data on corporate filings of new IPOs among VC-backed startups. Every firm filing an S-1 for an IPO must provide three years of compensation history for the CEO and other top executives. For each VC-backed startup that had an IPO in VentureSource, we find its S-1 filing (if available). We extract the executive compensation data from the filing and use a fuzzy string merge to flag the founder (if they are still with the firm). If the S-1 does not contain compensation data, then we collect the first post-IPO DEF 14-A proxy. For the data analysis, we collect these proxy statements and extracted the CEO compensation components for all firm-years. The CEO was identified in these filings using a fuzzy string search of titles in the filing and the compensation was scraped (hand-checked when inconsistencies emerged).

Next, we use Execucomp compensation data to assess the relationship between CEO compensation and firm size of public firms. With this data, we can compare this relationship to that found in AHR for private firms. We downloaded the Execucomp data for the universe of U.S.-based Compustat firms. Because Execucomp covers only a subset of public firms, we fill in gaps with the scraped DEF 14-A data described above. We merge the compensation data of the CEO to the annual revenue data reported in Compustat for U.S.-headquartered firms.

Figure A1: Differences in Residual Log Cash Compensation Between Founder CEOs and non-Founder CEOs in VC-backed Startups

Notes: This figure plots residuals from regressions of log total pay when different controls are included. Residuals come from pooled regressions (winsorized at the 1% level from below) and are then plotted separately for founder and non-founder CEOs. The "Base" panel (A) comes from a regression that includes year and industry fixed effects. Panel (B) adds revenue and headcount fixed effects. Panel (C) adds firm age and Panel (D) adds fixed effects for the amount of capital raised and the number of funding rounds.

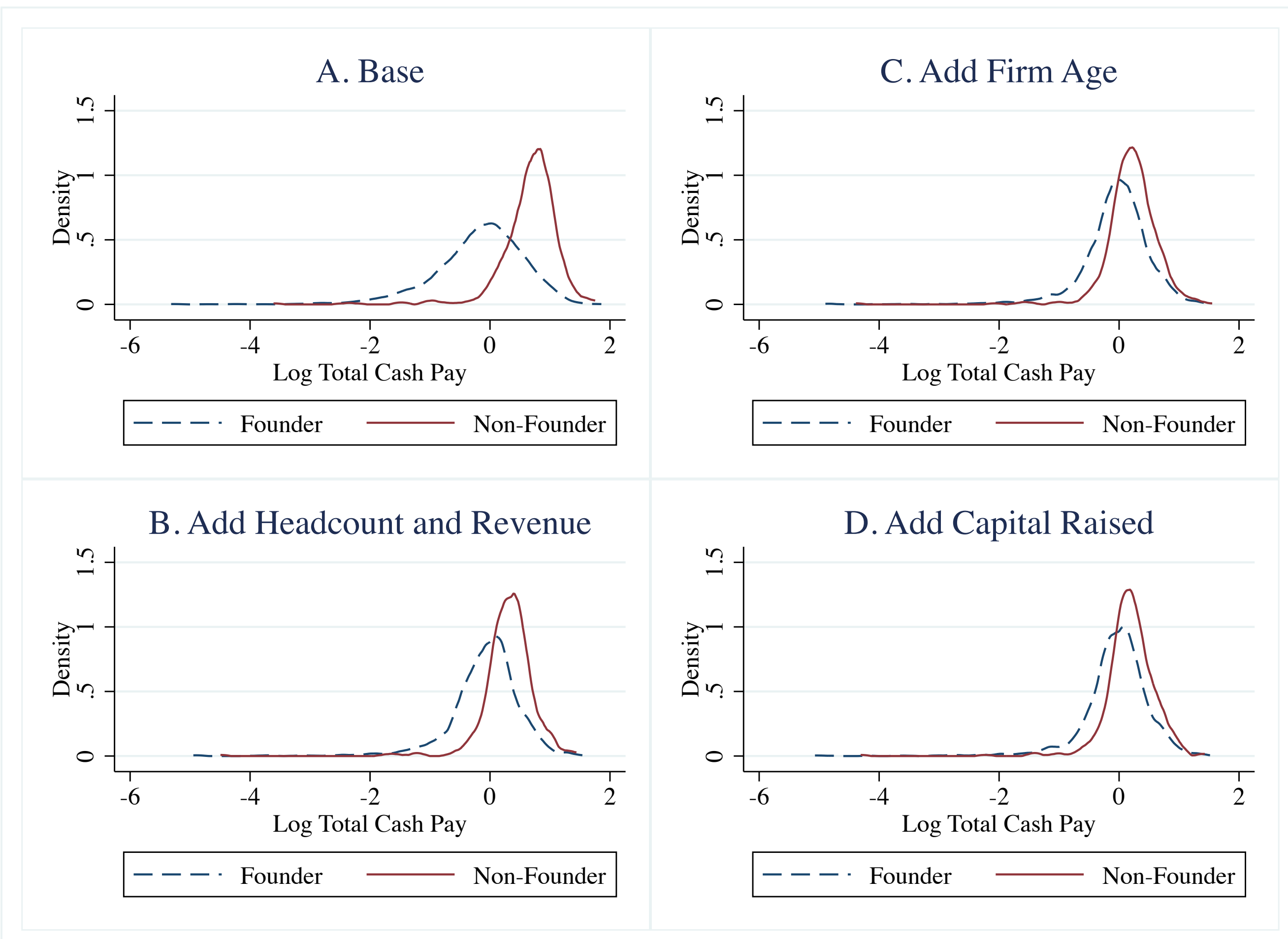


Table A1: Comparison of AHR Compensation Data and Venture Investments by Eligible Funds

Notes: This table compares the AHR data to the VC investments in VentureSource, conditioning on portfolio companies of the funds that participate in the AHR survey in each year. The VentureSource data were downloaded in the fourth quarter of 2018 and, to match the AHR sample, we exclude firms in healthcare, energy, industrial goods, and unknown industries. We exclude firms in both samples that are greater than 10 years old. We drop VentureSource firms that have more than 4 years elapsed since raising their last round of capital. We exclude non-standard financings (e.g., Initial Coin Offerings, spinoffs, or restarts) and also attempt to exclude mature firms that access private capital later in their lifecycle by dropping firms that have traditional private equity (rather than VC) investors in the first round, large debt rounds in their first round of financing (greater than \$3 million), or first rounds of financing with greater than \$50 million in investment.

	Venture Source Sample				AHR Sample				
	2015 Count	2017 Count	Pooled Count	Share of Pooled Venture Source Sample	2015 Count	2017 Count	Pooled Count	Share of Pooled AHR Sample	AHR Fraction of Venture Source
Panel A: By Firm Age									
Founded This Year	74	51	125	3%	29	49	78	3%	62%
1 Year Old	159	181	340	9%	108	161	269	12%	79%
2 Years Old	256	323	579	15%	133	215	348	15%	60%
3 Years Old	311	334	645	17%	133	203	336	15%	52%
4 Years Old	291	332	623	16%	119	210	329	15%	53%
5 Years Old	187	320	507	13%	63	185	248	11%	49%
6 Years Old	161	253	414	11%	72	136	208	9%	50%
7 Years Old	114	164	278	7%	58	80	138	6%	50%
8+ Years Old	150	238	388	10%	112	190	302	13%	78%
Total	1703	2196	3899	100%	827	1429	2256	100%	58%
Panel B: By Total Venture Capital Raised (Millions of Dollars)									
\$0 - \$10 Million Raised	649	929	1578	40%	320	628	948	42%	60%
10--25	419	464	883	23%	163	265	428	19%	48%
25-50	296	373	669	17%	151	245	396	18%	59%
50-75	125	176	301	8%	80	108	188	8%	62%
75-100	76	74	150	4%	45	69	114	5%	76%
100+	138	180	318	8%	68	114	182	8%	57%
Total	1703	2196	3899	100%	827	1429	2256	100%	58%

Table A2: Comparison of AHR Compensation Data and the Universe of Venture Investments

Notes: This table compares the AHR data to the universe of VC investment in VentureSource. The VentureSource data were downloaded in 2018 Q4 and, to match the AHR sample, we exclude firms in healthcare, energy, industrial goods, and unknown industries. We exclude firms in both samples that are greater than 10 years old. We drop VentureSource firms that have more than four years elapsed since raising their last round of capital. We exclude non-standard financings (e.g., Initial Coin Offerings, spinoffs, or restarts) and also attempt to exclude mature firms that access private capital later in their lifecycle by dropping firms that have traditional private equity (rather than VC) investors in the first round, large debt rounds in their first round of financing (greater than $3 million), or first rounds of financing with greater than $50 million in investment.

	Venture Source Sample				AHR Sample				
	2015 Count	2017 Count	Pooled Count	Share of Pooled Venture Source Sample	2015 Count	2017 Count	Pooled Count	Share of Pooled AHR Sample	AHR Fraction of Venture Source
Panel A: By Firm Age									
Founded This Year	230	133	363	4%	29	49	78	3%	21%
1 Year Old	573	418	991	11%	108	161	269	12%	27%
2 Years Old	757	691	1448	16%	133	215	348	15%	24%
3 Years Old	833	753	1586	18%	133	203	336	15%	21%
4 Years Old	692	727	1419	16%	119	210	329	15%	23%
5 Years Old	462	643	1105	12%	63	185	248	11%	22%
6 Years Old	315	503	818	9%	72	136	208	9%	25%
7 Years Old	236	334	570	6%	58	80	138	6%	24%
8+ Years Old	341	383	724	8%	112	190	302	13%	42%
Total	4439	4585	9024	100%	827	1429	2256	100%	25%
Panel B: By Total Venture Capital Raised (Millions of Dollars)									
$0 - $10 Million Raised	2851	2817	5668	63%	320	628	948	42%	17%
10--25	757	779	1536	17%	163	265	428	19%	28%
25-50	427	499	926	10%	151	245	396	18%	43%
50-75	158	208	366	4%	80	108	188	8%	51%
75-100	89	87	176	2%	45	69	114	5%	65%
100+	157	195	352	4%	68	114	182	8%	52%
Total	4439	4585	9024	100%	827	1429	2256	100%	25%

Table A3: Descriptive Statistics on Founder-CEO Compensation from 2008–2010 Compensation Pro Data

Notes: This table displays mean and median total Post-Product compensation for founder-CEOs and the number of observations by industry and product/firm status for 2008–2010. The data come from a survey of startups done by VentureSource Compensation Pro. This sample includes additional industry categories that are not available in the AHR sample. Note that the pre-product sample does not appear representative of the VentureSource Universe for this time period. In the VentureSource sample for 2009, 54% of firms are pre-product, where a 2.45% of Compensation Pro firms are listed as Pre-Product.

	Number of Observations	Mean Salary for CEO of Post-Product VC backed firm	Median Salary for CEO of Post-Product VC backed firm
Banking and Finance	102	264,806	225,000
BioTech	172	295,686	281,000
Consumer Goods	21	183,810	185,000
Consumer Services	45	249,701	220,000
Energy and Utilities	50	235,531	210,000
Non-BioTech Healthcare	36	275,325	236,250
Industrial Goods	43	219,447	210,000
Information Technology	247	240,789	225,000

Appendix B Calibration

B.1 Data Processing

This section provides additional detail for how we process the data for the calibration exercise. Refer to Section 3.3.1 in the text for the main details about the VentureSource Sample.

B.1.1 VentureSource Data

There are several data cleaning and imputation steps that we take to prepare the VentureSource main sample for our calibration exercise. Recall that the main sample includes firms founded between 1987 and 2006. We start with a full download of VentureSource from 2020Q1 that includes exit information for all startups financed since the start of the sample. We exclude from this any firms that only raised growth rounds, which in VentureSource are typically PE-backed firms. A startup is included in the sample if it raised at least one equity financing round from a venture capital investor between 1987 and 2006. We drop financing rounds with missing close dates and firms that only raised debt financing. The final data includes information on capital raised, product status (updated with each new financing event), members of the investor syndicates, valuations (often missing), industry classifications provided by VentureSource, and executive team information.

B.1.2 Computing Returns to Founders

There are several steps that we take to compute the value of the firm upon exit to common shareholders add the founder's share of any exit value, as the firm's exit value is distinct from the return accruing to the founder, X. There are two issues with mapping the observed total exit value to what the founder receives. i) No standard dataset contains the full cap table for any entrepreneur-firm pair. ii) The founders' equity vests and is subordinate to investors' because of liquidation preferences. Hall and Woodward deal with the first issue by constructing the expected distribution of founder equity holdings from the cap tables in S-1 filings. They then integrating over the distribution of founder equity when calculating the founders' portion of an exit. Because the AHR data contains equity holdings for firms at different points in the lifecycle, our approach is instead to merge average founder-CEO equity holdings at the firm age -by- coarsened cumulative capital raised categories into the VentureSource data. This also enables us to study non-founders at lower levels of the corporate hierarchy in Section B.4, as the AHR data contain their equity holdings whereas S-1 filings do not. We then adjust the equity holdings observed in AHR to account for time series differences in the funding environment and different industry conventions for founder holdings based on moments from the Compensation Pro data. To adjust the equity holdings for firms in cloud-affected industries founded prior to 2006, we take 92% of the baseline equity share in the AHR data, which is the ratio of the Compensation Pro holdings to the AHR holdings. For firms in non-cloud affected industries, we take 79% of the equity share reported in AHR.

To account for founders' subordinate positions due to liquidation preferences, we condition on the exit type and whether the exit returns more or less than the capital invested in the firm. For IPOs, we assume that all shares convert to common equity, but founder holdings are diluted by 10% in the IPO. For acquisitions that return less than the capital invested, investors are assumed to exercise their liquidation preferences, and founders receive no value from their equity. For acquisitions that are between one times and 2.5X the capital invested, we assume the total value to common stock holders is equal to the exit value less the capital raised. For acquisitions that exceed 2.5X capital invested, we assume all investors convert to common. For acquisitions that occur prior to standard vesting periods, we assume that founders' equity vesting accelerates by 1.5 years, up to their being fully vested. If an acquisition value is not observed, we apply these rules to imputed values. Our imputation procedure is detailed in Appendix Table B2.

B.2 Solving the Model

We solve the model first for the post-product stage. When we report separate results by industry group, we solve the model industry-by-industry, while our main specifications use exit moments that are specific to cloud affected and non-affected industries. Following Hall and Woodward, we use an exponential grid of 500 asset values and a fixed outside wage w^*. We compute exit-probabilities by age from the VentureSource data for each industry. We backsolve from the terminal period, computing the agent's consumption decision, which allows us to recover value functions for each of the 500 asset levels by firm age. This directly follows from Hall and Woodward's approach, with the one change that the wage during the startup period evolves over the startup lifecycle. To keep the problem simple, we avoid having $E(w_t|P)$ depend on w_{t-1}, which would make the wage an additional state variable. Instead, t and product status capture all decision-relevant information about the expected wage. This allows for stochastic wage realizations, while avoiding the need to build in the wage history as a state variable in the model. We use the asset-by-firm age value function array as an input into the pre-product program.

We use a similar approach to solve the pre-product program, but we build in a product transition probability and a pre-product exit probability as described in the text. These probabilities are are estimated from the VentureSource data from the empirical transition densities. In the pre-product program, we weight the asset-by-firm age array that captures the post-product value function by the product transition probability.

Table B1 shows the hazard of transitioning from the pre-product to post-product state for firms in Cloud affected and Non-Cloud affected industries before and after 2006. The pre-product to post-product transition process and the pre-product exit processes are estimated conditional on industry for the pre- and post-periods. In the post-period the hazard of transitioning to a product increases for 2, 3, 4, and 5 year old firms in cloud-affected industries. The hazard is flat for non-cloud affected industries across time despite the product transition hazard being always lower for non-cloud firms. Table B1 shows the hazard of exiting prior to achieving product status. For cloud-affected industries, the hazard in the post-period falls for firms under two years old and then increases for firms after year three. For non-cloud industries, the hazard falls in year two, is roughly constant in years three and four, and increases in year five.

Numbers reported in tables and figures are typically the minimum value of assets/wealth, given an outside salary w^*. We calculate these values as the minimum value in the asset grid for which the pre-product value function in (3) exceeds the value function from not entering entrepreneurship, which we characterize for each wealth-salary combination using equation (2).

B.3 Validation: Founder Career Trajectories

In this section we present evidence on the plausibility of our maintained assumptions for the modeling exercise. First, recall that our model assumes the founders' wage is stable. If founding a startup accelerates career trajectories, either through providing human capital or signaling leadership propensity, we would expect to see the post-founding wage grow faster than our assumptions for non-founders. On the other hand, models of specialization predict that founders will potentially accumulate human capital that is too general, reducing their subsequent wage if they return to specialized roles (Lazear 2004).

Using our LinkedIn matched data, analysis of post-founding jobs compared to pre-founding jobs suggests that founding experience does not substantially alter one's expected earnings profile given the ACS occupations that we can match to. Appendix Figure B1 shows that the modal occupation transition between for founders before and after founding is zero. While there is also some variance that may suggest founders must bear some job search risk post-founding, there does not appear to be bias with more weight above or below zero. We also note that many non-founders also bear job risk, and richer models may build in frictions that necessitate some level of precautionary savings in the post-exit wage. It is possible, however, that founders bear risk if a venture's failure is more likely to lead to labor market penalties. Recent field experimental evidence does find that former founders' resumes receive fewer callbacks in audit studies of job applications, but the effect is driven by successful founders (Botelho and Chang 2022).

B.4 Startup Joiners

Startup employees are also an important part of input into innovative companies. In this section we briefly comment on who is willing to supply labor to startups given data on equity holdings and compensation for non-founders. Our data use agreement with AHR precludes us from disclosing compensation moments for non-CEO employees, but we are able to use the cash and equity compensation data for these individuals as inputs into the consumption/savings models that we use to study the attractiveness of entry for founders. Table B3 presents results where we consider senior joiners in tech positions (e.g., CTO), finance positions (e.g., CFO), sales and marketing positions (e.g., CRO), and other management positions (e.g., COO). Panel A considers who would find it worthwhile to join a pre-product startup at year zero. The "N/A" cells for those with outside salaries above $300,000 in "Finance and Other Management Positions" indicate there is no wealth level such that the expected equity return can rationalize the compensation gap between the outside salary and pay at the startup. Those in tech positions with a $300,000 outside salary would require over $7 million in wealth to join. For most senior joiner positions, it is only a positive expected-value decision for non-founder employees to join early stage, pre-product startups if their outside pay is below $around $225,000. Notice that the wealth thresholds increase rapidly in some cases: going from $110,000 for a finance executive with an outside salary of $225,000 to having no possible wealth level that rationalizes entry at a $300,000 salary. This pattern suggests that joiners' expected equity value does relatively little to contribute to the value of startup employment.

Panel B considers joiners of post-product startups at year zero. The wealth thresholds required to rationalize labor supply fall dramatically for those with a $300,000 outside salary, but they remain extreme for those with outside salaries above $400,000. Panel C considers a post-product startup that is three years old, but as mentioned in the table note, we continue to use the equity holdings data for joiners of startups that are less than two years old. This likely overstates equity grants for later joiners and gives us an upper bound on the attractiveness of supplying labor for joiners. In this case, very wealthy individuals in tech positions with outside salaries at $400,000 may find it attractive to join a startup, but few with salaries greater than $400,000 would be willing to join the typical post-product three year old startup. There may be exceptions for some firms with especially strong traction, as a three year old company will have more data on its likely prospects than a new firm in the “differentiation” phase, but on the whole highly compensated workers are not likely to be flocking to join young companies.

Figure B1: Changes in Founder-CEO Average Occupational Earnings in ACS Data.

Notes: This figure takes matched LinkedIn - ACS data for founders' occupations before and after founding. For each occupation, we compute mean annual earnings in the ACS data and take the post-founding less pre-founding difference in occupational averages.

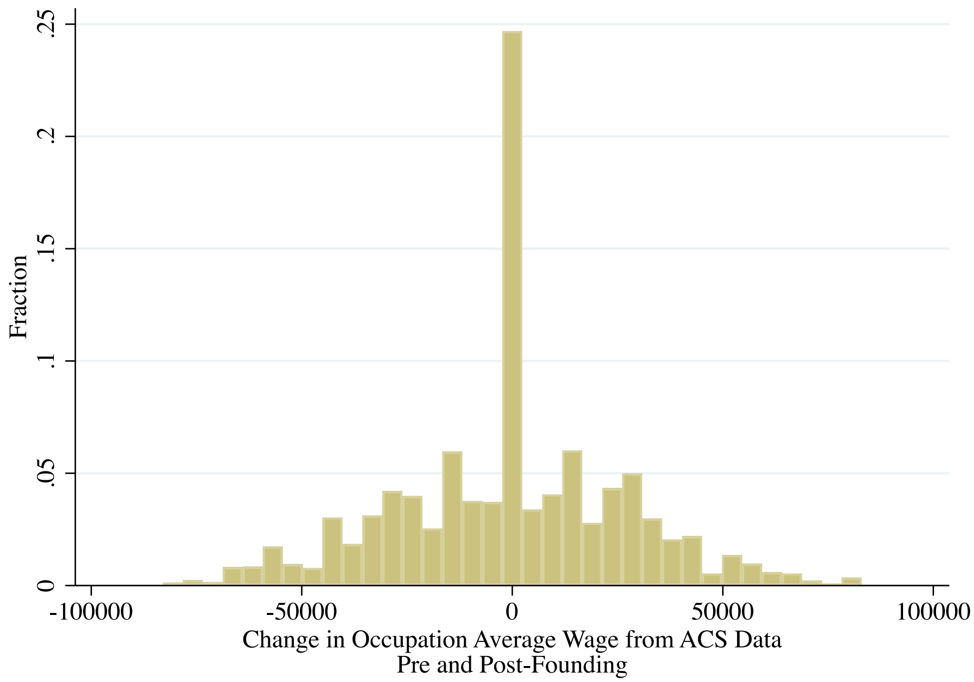
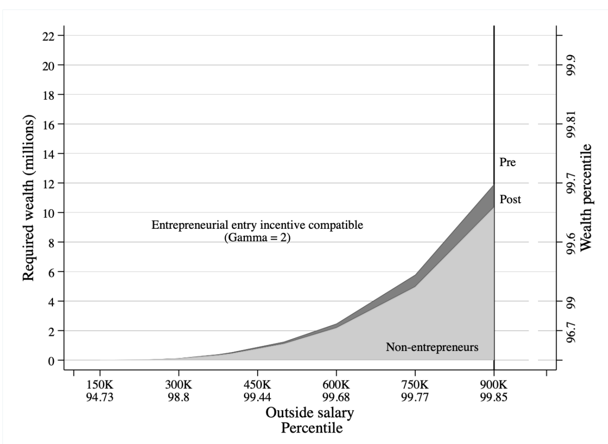
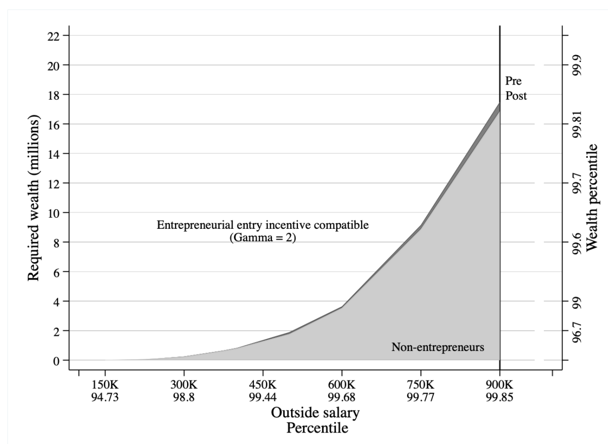


Figure B2: Heterogeneity in Time to Product and Attractiveness of VC-Backed Entrepreneurship When Holding Exit Values Fixed

Notes: This figure presents salary-wealth thresholds such that a potential entrepreneur will find founding attractive. It is the analog of Figure 5 from the main text, but rather than using the distribution of exit values by industry, we use a common set of exit values and vary only the product transition probability between the pre- and post-periods in cloud affected and non-cloud affected industries. That is, in each panel, we use the full sample of exits for the pre-period. We then vary the product transition and pre-product exit transition process as estimated by industry for the pre- and post-periods. These transition processes are given in Appendix Table B1. For these calibrations, we use compensation contracts that are product- and age-dependent and a coefficient of relative risk aversion of two.



(a) Cloud Product Transitions, Full Sample



(b) Non-Cloud Transitions, Full Exit Sample

Table B1: Product And Exit Transition Probabilities Used in Calibration

Notes: This table presents the estimated product transition and pre-product exit probabilities as estimated from the VentureSource data that we use to calibrate the consumption-savings models. These are probabilities of an event conditional on survival to firm age t.

	Cloud Affected Industries				Non-Cloud Affected Industries			
Firm Age	Post-Product Transition Probability		Pre_Product Exit Probability		Post-Product Transition Probability		Pre_Product Exit Probability	
	Pre-Period	Post-Period	Pre-Period	Post-Period	Pre-Period	Post-Period	Pre-Period	Post-Period
1	0.207	0.204	0.031	0.022	0.110	0.122	0.020	0.023
2	0.263	0.301	0.047	0.038	0.142	0.135	0.040	0.025
3	0.300	0.370	0.049	0.055	0.163	0.175	0.054	0.053
4	0.274	0.339	0.032	0.065	0.154	0.163	0.068	0.059
5	0.257	0.346	0.031	0.077	0.149	0.168	0.071	0.093
6	0.239	0.271	0.028	0.208	0.161	0.131	0.073	0.182
7	0.298	0.298	0.022	0.022	0.171	0.171	0.095	0.095
8	0.360	0.360	0.032	0.032	0.171	0.171	0.111	0.111
9	0.492	0.492	0.027	0.027	0.198	0.198	0.125	0.125
10	0.849	0.849	0.063	0.063	0.241	0.241	0.162	0.162

Table B2: Details about the VentureSource Sample Used in the Model

Notes: This table reports details of the VentureSource sample used to estimate firm transitions and exit values. The first row is a count of firms observed in VentureSource based on the founding date windows in each column. Row 2 displays the total number of firms in cloud affected industries, as defined by Ewens, Nanda, and Rhodes-Kropf (2018). These industries are “Business Support Services,” “Consumer Information Services,” “Financial Institutions and Services,” “Media and Content,” “Medical Software and Information Services,” “Retailers,” “Software,” and “Travel and Leisure.” Row 3 displays the total number of observed exits by the end of the sample, including firms that have missing exit values. This is the number of firms that VentureSource does not list as remaining private at the end of the sample. Exit values are not observed for all firms that VentureSource classifies as having exited. Row 4 gives the number of missing exit values among firms that have exited. Row 5 displays the number of firms for which we can impute exit values among those with missing values. We do this imputation by taking a random draw of the exit to common shareholders for firms in the same founding cohort, in the same quartile of age at exit, in the same quintile of total venture capital raised, with the same exit type (acquisition or IPO), with the same product status, and in the same cloud/non-cloud group. Row 6 indicates the number of pre-product exits with missing values that we set to zero. Row 7 indicates the number of firms that fail (with status either “Out of Business” or an exit value below total capital invested) for which we can observe the failure date. Row 8 shows the number of firms for which we impute failure times if they are unobserved. This imputation is done taking a draw from a beta(2,5) distribution on the 0-3 interval and adding 2 years to account for the average time between funding rounds. Because our dynamic program extends through 17 years post founding, row 9 shows the number of firms that remain private for which we impute exit values. Row 10 shows the number of firms with non-missing product dates, recorded at the next funding round after achieving a product transition. Row 11 shows the number of firms with product but missing product dates. Row 12 shows the number of firms for which we impute missing product dates based on whether the firm was founded before or after 2007, whether it is in a cloud-affected industry, its age at exit, and the delay between founding and the first VC funding round.

		Firms Founded Before 2007	Firms Founded Between 2007 and 2014
1	Total Firms	18,829	15,032
2	Total in Cloud Affected Industries	11,006	10,381
3	Firms Known to Have Exited (Not Private at End of Sample)	17,518	7,718
4	Firms from Row (3) with Missing Exit Values	4,113	2,074
5	With Exit Values Successfully Imputed By Year Founded, VC Raised, Exit Type, Product, and Cloud	4,061	2,060
6	With Non-Imputed Pre-Product Acquisitions Set to Zero	34	12
7	Firms that Fail (0 Exit to Common) with Failure Times Observed	6,385	1,449
8	Firms with Failure Times Calculated by Time Between Rounds or Beta Distribution over 5 Years	2,354	2,901
9	Firms that Remained Private through 2021 with Imputed Exit Times and Values Based on VC Raised and Cloud	1,309	N/A
10	Firms with Non-Missing Product Dates if Product at Exit or End of Sample	12,619	10,786
11	Firms with Missing Product Dates But Product at Exit	3,434	1,614
12	Firms with Successfully Imputed Product Dates by Cloud/Non, Age at Exit, and Delay Between Founding and First VC	3,432	1,614

Table B3: Analysis of Outside Salary and Wealth Combinations to Rationalize Joining a Startup in a non-Founder Role

Notes: This table shows conditions by outside salary and wealth where it is incentive compatible to join a pre-product startup, a post-product startup in year 0, or a post-product startup in year 3 as a senior non-founder (e.g. CTO or COO) in a functional area. The calibration exercise uses the pooled VentureSource sample of companies founded prior to 2006. Pay data and fully diluted equity holdings are taken from AHR. We compute exits to joiners by adjusting exits to founder-CEOs by the ratio of fully diluted equity held by joiners for firms under two years old to the ratio of founder-CEO equity. We use this measure of the ratio for young firms because our thought experiment considers joiners at the early stage, while the data on fully diluted equity at later stages may reflect that joiners who do not join pre-product firms may receive lower amounts of equity. The cells with "N/A" indicate that there is no incentive compatible wealth level for a given salary outside salary that would justify joining a startup. This is because the expected gap in pay over the life of the startup is greater than the expected return from an exit. Differences in wealth thresholds across columns for the same outside salary can arise because expected flow compensation varies by position or because fully diluted equity holdings differ by position. Although we cannot disclose equity holdings summary statistics, we note that non-founders in tech positions on average hold the most equity and non-founders in sales/marketing positions hold the least. Sales and Marketing positions have the highest cash compensation at both the pre-product and post-product phase. This likely arises because of commission based pay as part of the compensation contract.

	Tech Positions	Finance Positions	Sales/Marketing Positions	Other Mangement Positions
Outside Salary (100k)	Panel A: Wealth Required (Millions) for Joining an Early Stage Startup at Year 0 to Have Positive Expected Utility			
150	0.02	0.02	0.02	0.02
225	0.08	0.11	0.06	0.11
300	6.86	NA	3.92	NA
400	N/A	N/A	N/A	N/A
Outside Salary (100k)	Panel B: Wealth Required (Millions) for Joining a Post-Product Startup at Year 0 to Have Positive Expected Utility			
150	0.02	0.02	0.02	0.02
225	0.05	0.06	0.05	0.05
300	0.29	3.34	0.21	0.29
400	N/A	N/A	N/A	N/A
Outside Salary (100k)	Panel C: Wealth Required (Millions) for Joining a Post-Product Startup at Year 3 to Have Positive Expected Utility			
150	0.02	0.02	0.02	0.02
225	0.04	0.05	0.04	0.04
300	0.10	0.13	0.09	0.09
400	14.92	N/A	N/A	N/A

Appendix C Empirical Analysis of Founder Backgrounds and Selection

C.1 Identifying Founder-CEOs in the LinkedIn Data

The first step in linking our LinkedIn and VentureSource data is finding a startup that appears in the VentureSource data in LinkedIn profiles. We initially use the startup's name and website URL to find the startup's unique LinkedIn profile. We validate the match by comparing the company name in VentureSource to that in the LinkedIn profile. We then search among all employees of this firm in LinkedIn with "founder" or "CEO" in their title. Because some founders will only list their title as CEO on LinkedIn, we consider only CEOs to be founders if they join the company prior to 90 days since raising their first VC financing. If there are multiple individuals that have both founder and CEO in their title, we select the one with the earlier of the join dates. Individuals without the startup in their LinkedIn profile are dropped. Our final dataset has 9,574 founder-CEOs over the 2001 to 2016 period.

C.2 Representativeness of the LinkedIn Founder-CEOs and Differential Changes Over Time

We next compare the set of firms where we can find founder-CEOs' LinkedIn profiles to those that are unmatched. Table C1 presents summary statistics for a set of startup observables in the 2001–2016 sample period split by whether we find a founder-CEO in LinkedIn. Those startups with a LinkedIn founder ("With match to founder-CEO") are different on several dimensions including first capital raised, exit valuations, and share with a product by startup exit. Moreover, we see that the industry compositions differ: matched firms are more likely to be information technology firms, which include cloud affected industries. The differences in capital raised are likely explained by the differences in the number of software and IT firms, which require less capital. The higher exit valuation and share with a product likely stem from survivorship and firm size: startups and their employees are more likely to have profiles when there is a significant amount of hiring or employee turnover. Both are more likely to occur as the firm raises more capital. The higher coverage for cloud-affected industries leads us to re-weight the LinkedIn sample to reflect the count of VentureSource firms by detailed VentureSource industry and founding year.

Although we find some differences between firms that we can match to LinkedIn founders and those that we do not, any bias in our matching process is only relevant if the matching systematically differs across industries and time cells in a way that would affect our inference about changes in speed to product for cloud industries. Table C2 runs a triple differences test of this possibility. We examine the size of the first round of financing and total financing raised (as proxies for unobserved founder quality when we cannot find a LinkedIn profile). We interact indicators for firms where we find a startup founder in LinkedIn "LinkedIn," with post-period indicators and indicators for cloud affected industries. Founders in LinkedIn do appear to raise more capital at baseline in the pre-period, with first round sizes that are about 27-28% larger than those for firms' founders we do not observe. However, there are no differences in relative round sizes for founders in cloud affected industries versus non-cloud industries in the per-period, which suggests our difference-in-differences estimates are not capturing differential bias in firm or founder quality between cloud and non-cloud industries. LinkedIn coverage for firms founded across all industries improves in the post period, as the "Post X LinkedIn" indicator offsets the pre-period effect of LinkedIn covering larger firms. If anything, this tells us that LinkedIn captures more marginal startups in the post-period. Consistent with the notion that the cloud shock reduces the need to raise financing, cloud affected firms in the post-period raise smaller initial rounds. These double differences are indicative of patterns across industries and time, but they are not informative about bias for our difference-in-differences analysis.

Our main test is the triple differences indicator of "Post x Cloud x LinkedIn," which captures whether the average quality of firms (proxying for founders) in LinkedIn changes differentially across industries. The answer is no. These coefficients are never statistically different from zero, suggesting any bias in our LinkedIn match is not correlated with the decline in time to product in cloud affected industries.

C.3 Matching to the ACS

Matching to the ACS data requires us to know the NAICS industry codes and Standard Occupation Classification (SOC) codes associated with LinkedIn firms and occupations.[27] We use a token-based fuzzy matching algorithm to link firm names in LinkedIn with firm names in Burning Glass, which contains firm-to-NAICS code mappings.[28] For firms that we cannot match to a Burning Glass NAICS code, we use LinkedIn's industry keywords and take the modal NAICS code of firms that

[27]In about 6% of cases, the founder-CEOs pre-startup work experience is outside the U.S. The ACS data covers only U.S. firms, so we use the LinkedIn company's industry and founder-CEO's occupation title to match to ACS.

[28]Burning Glass is a data provider that scrapes job ads. For every job ad, Burning Glass fills in NAICS and SOC codes based on the firm and job title.

we can match to Burning Glass data. We use a similar procedure to get SOC codes, where we match job titles to the job titles advertised in Burning Glass.[29]

[29] The text of job titles is less standardized than firm names. To augment our matching procedure, we rely on the Department of Labor's O-NET database, which contains a mapping of many "Alternative Job Titles" to SOC codes. This data is used for many administrative purposes to map the raw text of job names to the SOC classification scheme. When there is disagreement between the match between sources, we consider a "high" and "low" version based on which occupation has higher earnings in the ACS data.

Table C1: Comparing Startups in VentureSource with a Match to LinkedIn Founder-CEO to Those Without

Notes: This table reports the differences between the full sample of startups founded between 2001 and 2016 to those with an identified LinkedIn founder-CEO and without such a match. "Amount raised" is the amount of capital the startup raised in its first financing. "Share with IPO" is the percent of firms with IPO exits by the end of the sample. "Estimated Exit Value" is the exit value (if known) by end of sample. "Year first VC" is the year of first VC financing. "Share that got to post-product before exit" is the share of firms that had a product by end of sample or their exit. The remaining variables report the share of firms in each major VentureSource industry.

	Venture Source Sample 2001-2016	With match to founder-CEO	Without match to founder CEO
Number of Startups	26,102	9,574	16,528
Share of Startups		37%	63%
Amount Raised	6.2	5.4	6.7
Share with IPO	0.02	0.02	0.02
Exit Valuation	106.7	156.0	82.6
Year first VC	2012	2012	2012
Share that got to post-product before exit	73%	79%	69%
Information Technology	32%	34%	31%
Healthcare	20%	16%	22%
Business/Consumer/Financial	43%	46%	41%
Other industry	6%	5%	6%

Table C2: LinkedIn Representativeness: Difference-in-Differences

Notes: The table reports triple differences regressions of a startup's first round of capital raised and total capital raised (as of the end of the sample). It compares startups in the main sample that have a LinkedIn founder-CEO match ("Linkedin"), in cloud industries ("Cloud") and founded in the post-2006 period ("Post"). Columns 1-2 have the log of first capital raised as the dependent variable and 3-4 the log of total capital raised. The even columns use weights to match the LinkedIn matched sample to the industry-founding-year distribution in the full sample. All regressions include year fixed effects and robust standard errors reported in parentheses.

	First Capital Raised		Total Capital Raised	
	(1)	(2)	(3)	(4)
Post X Cloud X LinkedIn	0.130	0.069	0.037	-0.054
	(0.010)	(0.11)	(0.120)	(0.130)
Post X Cloud	-0.290***	-0.290***	-0.044	-0.040
	(0.056)	(0.056)	(0.065)	(0.065)
Cloud X LinkedIn	-0.066	-0.058	-0.043	0.003
	(0.081)	(0.084)	(0.099)	(0.100)
Post X LinkedIn	-0.230***	-0.210**	-0.270***	-0.230**
	(0.083)	(0.089)	(0.097)	(0.10)
Cloud-affected industries	-0.100**	-0.100**	-0.320***	-0.320***
	(0.043)	(0.043)	(0.051)	(0.052)
In LinkedIn	0.270***	0.280***	0.630***	0.600***
	(0.066)	(0.069)	(0.079)	(0.084)
Weighted	No	Yes	No	Yes
R-Squared	0.044	0.057	0.068	0.068
Observations	22,425	22,425	23,581	23,581