ABSTRACT

Title of Dissertation: Essays on Entrepreneurship: The Role of Complexity of Innovation and Efficient Hierarchies

Yuheng Ding, Doctor of Philosophy, 2023

Dissertation directed by: Professor Serguey Braguinsky, Ph.D.
Department of Management and Organization
Robert H. Smith School of Business

Entrepreneurial activities have been on the decline across a broad range of sectors in the U.S. during the past few decades. This decline (sometimes also called "declining business dynamism") is reflected in the decreasing rate of new firm entry, the share of young firms (usually defined as those five years of age or less) in the total number of firms and/or the share of employment at young firms in total employment, and so on (e.g., Decker et al. 2014; Akcigit and Ates, 2021). All of the above have exhibited a secular decline, not just in the U.S. but in other advanced economies as well. The underlying causes of these trends, however, are not yet clear with a broad array of explanations suggested in the literature (Akcigit and Ates, 2019; 2021; Decker et al., 2016; Hopenhayn et al., 2018; Karahan et al., 2019; Andrews et al., 2016). There also appears to be a lot of heterogeneity in how strongly the decline in entrepreneurial activities (business dynamism) is pronounced in various industries and sectors of the economy. In particular, the evidence in Haltiwanger et al. (2014) suggests that high-tech industries could be affected more than other sectors of the economy. High-tech sectors have been the driving force of growth in

recent decades, so uncovering the reasons for declining business dynamism in those sectors is a task of first-order importance.

In the first chapter, I employ the restricted-use data on the science and engineering workforce in the U.S. to investigate whether the increasing burden of knowledge is a growing concern for science-based entrepreneurship. Results show that since 1997, the rate of startup formation has precipitously declined for firms operated by U.S. Ph.D. recipients in science and engineering. The decline in startup formation is accompanied by an earnings decline, increasing work complexity in R&D, and more administrative work for science-based founders. With limited access to efficient knowledge hierarchies, founders of science-based startups must shoulder the burden of knowledge by doing more tasks by themselves. Workers at established firms, on the other hand, could better mitigate the burden of knowledge by narrowing the span of control and increasing the depth of hierarchy. Moreover, less experienced founders were hit harder than more experienced founders as the increasing burden of knowledge led to increasing returns to labor experience.

While in the first chapter I use individual-level work data, in the second chapter I utilize firm-level data from the U.S. Census Bureau to develop the analysis further. I adopt the abductive approach and leverage matched employee-employer Census data between 2000-2014 to investigate how a growing burden of knowledge (measured as knowledge interdependence) in the most innovative firms affects potential entrepreneurs' decisions to start their own business ventures. I show that higher knowledge interdependence in incumbent firms is negatively associated with employee entrepreneurship, and the negative effect is pronounced even stronger among the highest-performing employees. Moreover, higher knowledge interdependence has a positive selection effect on the quality of "spinouts", and this effect is significantly stronger if the startup is formed by individuals ranked highest in the human capital distribution. These results suggest that knowledge interdependence does not merely raise the barrier for entry into entrepreneurship by imposing higher costs of knowledge transfer. It also changes the functioning of the internal labor market inside the firms.

In the third chapter, I further investigate the mechanism underlying the relationship between knowledge interdependence and employee entrepreneurship. I propose a formal theoretical framework that reconciles all empirical findings. The theory suggests firms that rely on higher knowledge interdependence should share "rent" with their employees by paying wage premia if the profit from higher knowledge interdependence is high enough. As a result, within-firm earning dispersion would always be larger in firms relying on higher knowledge interdependence. I find supporting evidence in the data for this alternative explanation.

Overall, these findings have important implications for declining entrepreneurial activity, rising income inequality, and technological change in the U.S. economy. While the conventional wisdom might view the declining entrepreneurial activity in the U.S. as the demise of economic growth, it is possible that as innovation becomes more complex, large established firms start to substitute the role of start-ups in pushing forward the technological frontier and driving economic growth as the efficient knowledge hierarchy could better deal with complex knowledge needed in the production process (Garicano, 2000; Garicano and Rossi-Hansberg, 2004). If this is the case, the declining business dynamism might just be a reflection of technological change and efficient (re)allocation of resources but not necessarily detrimental to technological advancement and economic growth. Whether this is true remains an avenue for future research.

ESSAYS ON ENTREPRENEURSHIP:

THE ROLE OF COMPLEXITY OF INNOVATION AND EFFICIENT HIERARCHIES

by

Yuheng Ding

Dissertation submitted to the Faculty of the Graduate School of the
University of Maryland, College Park, in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
2023

Advisory Committee:
Professor Serguey Braguinsky, Chair
Professor Rajshree Agarwal, Co-Chair
Professor Brent Goldfarb
Professor Evan Starr
Professor Martin Ganco
Professor David Sicilia

Dedication

I dedicate this dissertation to my grandparents, Xinghua Zhang (张兴华) and Qingping Wang (王青萍), for without their early inspiration and example of unyielding integrity, insatiable curiosity, and unwavering perseverance, none of this would have happened.

Acknowledgements

Words cannot express how grateful I am to Serguey Braguinsky, my advisor, mentor, and friend, for his invaluable guidance, support, and encouragement. He is a scholar in the truest sense, and his dedication to scholarship and intellectual curiosity have been contagious. He always inspires me to tackle big issues faced by our society, to challenge convention, and to pursue my passions with confidence and conviction. He is not only the best academic father I could ever possibly have, but also a caring friend who has always been there for me as a person and as a scholar.

I would also like to express my deepest appreciation to my co-advisor Rajshree Agarwal, for being extremely considerate, incredibly inspiring, and her unreserved support including those I didn't even think of. Her advice and inspiration are invaluable for me to build my own human enterprise to pursue both excellence in scholarship and fearless entrepreneurship. Thank you, Serguey and Rajshree, for making my journey the destination.

I owe an incredible amount of thanks to my dissertation committee members. Brent Goldfarb has made me think about different types of inferences in philosophy seriously and completely changed my approach to research. I will always be channeling Evan Starr when reading regression tables and ask myself, "Are these omitted variables or bad controls? How do you interpret the coefficient estimates?" Their research method class is truly the best method class ever, really. I am so thankful to have Martin Ganco who gave me invaluable advice throughout the development of this dissertation and encouraged me to search for fundamental drivers of rising knowledge complexity. David Sicilia is truly a generous and thoughtful scholar, who inspired me to continue my inquiry to fully understand the social impact and implications on innovation of technological change.

Thomas Åstebro deserves special thanks for being an amazing co-author and senior colleague. He has taught me a great deal about the art and science of writing and publishing academic papers, from developing research ideas to crafting compelling arguments, and from navigating the peer review process to dealing with editorial feedback.

Each of the other faculty at the University of Maryland provided me priceless guidance and support during my journey here. Liu Yang has been my role model since I first transit from China to the U.S. in 2014, and she is the one who inspired me to pursue an academic career. Rachelle Sampson provided me with the very first chance to be involved in academic research. Waverly Ding gave me invaluable advice for the development of early-stage research ideas and tips on my job market presentation. David Kirsch nudged me to think through the most fundamental and philosophical questions about organizations, the nature of technology, entrepreneurship, and so on. David Waguespack has left an imprint on me that will not rub off anytime soon – "What is the difference between this identification strategy and the ideal experiment?"

I would also like to give thanks to my colleagues and friends at the University of Maryland: Ali Alghareeb, Nathan Barrymore, Mo Chen, Anthony Gibbs, Jay Habegger, Takuya Hiraiwa, Roxanne Jaffe, Seojin Kim, Benjamin King, Najoung Lim, Yongyi Lin, Sandeep Pillai, Runtian Tang, Marta Villamor, James White, Audra

Wormald, Beril Yalcinkaya, Shotaro Yamaguchi, and Liyue Yan for their collegial support and constructive feedback, comments, and suggestions. Jing Xue and Seula Kim deserve special thanks for being great work companions at the Census RDC. Justina Blanco's expertise, dedicated efforts, and genuine care make my time in the PhD program joyful.

Finally, my family deserves the biggest thanks of all. My mother, Jianwei Zhang (张建伟), has always been my pillar of strength. Her wisdom, compassion, and unconditional love has taught me the importance of caring for others and making a positive impact in the world. My father, Hua Ding (丁华), is a truly entrepreneurial role model in my life. He always inspires me to pursue my own entrepreneurial aspirations with fearless risk-taking, tireless work ethic, and positive attitude to overcome setbacks and challenges. My wife, Xinyue Zhang (张馨月), is my rock, my confidant, and my best friend. I could not have done this without her sacrifices, understanding, and unwavering belief in me even when I didn't believe in myself. To my 9-month-old son, Ethan, your innocent smiles, boundless curiosity, and playful energy have reminded me of the wonder of life and the importance of my research and the impact that it could have on shaping the future. This achievement is as much yours as it is mine, and I dedicate it to you all with my love and appreciation.

Table of Contents

Chapter 1: Declining Science-Based Startups: Strategic Human Capital and the Value of Working in Startups vs Established Firms[1]

Introduction

There is a deep concern that the U.S. has experienced a declining trend in entrepreneurship, as manifested in lower firm entry rates, falling employment share in new firms, and declining job mobility (e.g., Decker, Haltiwanger, Jarmin & Miranda, 2014; Goldschlag and Miranda, 2020). The reasons for these trends are still being debated.

In this article, we employ the nationally representative Survey of Doctorate Recipients (SDR) to show a decline over the past 20 years in both the rate of startups founded and the share of employment at startups by the highest-educated science and engineering portion of the U.S. workforce. The declines are wide-ranging and not driven by any particular demographic category or geographic region or scientific discipline.

Meanwhile, the burden / complexity of knowledge necessary for innovation has been on the increase (Bloom et al., 2020), as is the associated use and size of teams in science and invention (Jones, 2009; Wuchty, Jones & Uzzi, 2007). We conjecture that as the factor particularly relevant to science-based startups, this could be as a potential source of the afore-mentioned decline. As an example from medicine, "*It is estimated that the doubling time of medical knowledge in 1950 was 50 years; in 1980, 7 years; and in 2010, 3.5 years… What was learned in the first 3 years of medical school will*

[1] This chapter is co-authored with Thomas Åstebro and Serguey Braguinsky.

be just 6% of what is known at the end of the decade from 2010 to 2020" (Densen, 2011).

We employ a question-driven approach to explore how the increasing complexity of knowledge may negatively impact science-based founders, and therefore result in a decline in such startups. While startup founders as repositories of knowledge have been extensively examined (see literature review), the links between the troika of increasing complexity of knowledge over time, declining business dynamics, and startup founders' incentives and opportunities to be effective have not been clarified. The rich data available in SDR surveys provides us with an opportunity to explore these connections.

We use detailed individual-level work data from these surveys to document changes in the allocation of human capital in the highest-educated, science-based part of the U.S. workforce. In particular, we demonstrate that the differences in the nature of work, compensation, and allocation of talent have all changed dramatically between startups and large firms over time, and that these changes have over time increasingly favored working in established, and especially large firms, over startups. It is the differential ability of these two types of firms to rearrange their internal complementary assets to human capital that we think could be an important part of the increasing advantage of working in large firms over time, at the expense of declining startup creation. To the best of our knowledge, this study represents a first attempt to use rich individual-level data to examine how business dynamics in the knowledge-intensive and high-tech sectors may be affected by the changing nature of work.[2]

We document a decline of 38 percent in the rate of startup formation and a similar decline in the share of employment at startups by U.S. PhD degree holders in science and engineering over the past 20 years. The declines are universal, observed across all demographic categories, geographic regions, occupations, and scientific disciplines in the U.S.

We then turn to abductive reasoning to search for a plausible explanation of this trend.[3] For example, we show (a) a growing importance of work experience before becoming a founder; (b) an increase in the number of tasks that founders have to cope with; and (c) limitations faced by startups in utilizing efficient knowledge hierarchies. Workers in established firms have not experienced nearly as much increase in tasks; and have instead experienced increasing relative wages. Other trends that we document in the organization of work imply that large firms have been better at adjusting to the increase in the complexity of innovation than startups. Hence, an increasing complexity

---

[2] Most research examining the decline in business dynamics has focused on an economy-wide, "macro" perspective, with potential explanations centering around increased market concentration and the dominance of "superstar" firms (De Loecker, Eeckhout & Unger, 2020; Autor, Dorn, Katz, Patterson & Van Reenen., 2020), changing demographics (aging) of the U.S. population (Hopenhayn, Neira & Singhania, 2018; Karahan, Pugsley & Sahin., 2019), and so on. See Akcigit and Ates, (2021), for a recent review of this literature.

[3] While long recognized as one of the three basic forms of human reasoning (Pierce, 1997), the abductive reasoning is recently being rediscovered in the strategy, organization as well as economic literature (Pillai, Goldfarb, and Kirsch, 2021; Shani, Coghlan, and Alexander, 2020; Heckman and Singer, 2017; Tohmé and Crespo, 2013). Abductive reasoning "produces exploratory hypotheses" and "yields plausible explanations about puzzling phenomena … following a question, such as 'What is going on?'" (Shani, Coghlan, and Alexander, 2020, pp. 64-65).

of innovation appear to reduce the relative perceived value of founding new science-based firms.

Why should one care that innovative work has gotten, relatively, more comfortable in large firms? Not at all if one favors the concentration of innovative work to the already largest tech firms. But, if it is new firms from which many new technologies and business opportunities typically originate (Schumpeter, 1942) so that "[a] country cannot be great over a sustained period without a steady flow of great new firms" (Klepper, 2016, p. 62), then these trends are of great concern. Concluding the article, we discuss various potential actions that could be taken by policy makers, universities, founders and other start-up ecosystem builders to mitigate these trends.

The rest of the article is organized as follows. In Section 2 we provide a brief literature review. Section 3 describe the data, with more exhaustive information provided in the Appendix. In Section 4 we present our key findings. Section 5 examines various trends in more detail. Section 6 focus on the human capital of startup founders and their job satisfaction. Section 7 concludes.

Literature Review

The strategy literature has recognized human capital as a key source of firm competitive advantage (Campbell, Ganco, Franco, & Agarwal, 2012 Castanias & Helfat, 1991, 2001; Agarwal, Braguinsky, & Ohyama, 2020). This is especially true for startups (Carnahan, 2017; Klepper, 2016). As greater complexity of innovation causes greater coordination challenges (Rivkin, 2000; Nickerson and Zenger, 2004), a firm's capability to organize larger teams may become a crucial determinant for its success (Jones, 2009; Ganco, 2013; Agarwal et al., 2016). Work practices in science jobs have indeed been changing over time. For example, Jones (2009), and Wuchty et al. (2007) document an increasing preponderance of teamwork in science and invention.

Firms can implement a variety of human capital management practices to retain valuable human assets (see e.g., Campbell et al., 2012; Carnahan et al., 2012; Gambardella, Panico & Valentini, 2015; Wang, Zhao & Chen, 2017; Zenger, 1992). But if all firms follow these management practices, then nobody receives any competitive benefits in a general equilibrium. In our search for the best explanation for the observed relative decline in science-based startups, we explore how startups and large firms appear to *differ* in their ability to implement various human capital management practices to mitigate the impact of increasing knowledge complexity.

The literature on efficient knowledge hierarchies (Caliendo and Rossi-Hansberg, 2012; Garicano, 2000; Garicano and Rossi-Hansberg, 2004), can explain how large organizations, compared with startups, cope differentially with the increasing burden of knowledge. As the complexity of innovation increases, more innovation work exceptions are reported by lower-level workers to their managers, and ultimately to the C-suite. The firm can respond by either training workers to handle more of these exceptions or by increasing the number of layers to handle the increasing volume of reports, and by narrowing the variance in reports to allow greater job specialization. Indeed, Caliendo and Rossi-Hansberg (2012) show that larger firms are more likely to become hierarchically taller. Founders at startups should logically have less recourse to this mitigation strategy as they are typically financially constrained

(Evans and Jovanovic, 1989) and hence, less likely to have the financial resources to pay for either additional layers of hierarchy or to pay for training.

In sum, this brief literature review suggests that we should explore the differential evolution of a number of human resource practices, including compensation, task allocation, and hierarchical design, to enlist evidence of why science-based startups have been falling in relative numbers.

Data

The Survey of Doctorate Recipients (SDR) is a nationally representative survey of doctorate recipients in science and engineering from U.S. PhD-granting institutions. It is sponsored by the National Center of Science and Engineering Statistics (NCSES). WeI employ restricted-use data from all available surveys conducted (for the most part) biannually between 1995 and 2017 and report the results using NCSES-provided weights. We limit the sample to individuals residing in the U.S., employed full time (working at least 48 weeks per year and at least 30 hours per week) with non-zero salaries, who either work at their own business, or whose principal employer is a private-sector for-profit company at the time of the survey.

Information about startups comes from answers to the question whether one's principal employer "c[a]me into being as a new business within the past five years." We call businesses that did not come into being within the prior five years "established firms." We split the data to show the fraction of the startups in the total number of all firms (the sum of startups and established firms). This is consistent with the commonly used way of measuring the fraction of new firms and their employment share in the economy (see, e.g., Decker et al., 2014). Note that this concept of a startup rate is theoretically and empirically distinct from the concept of the self-employment rate (the fraction of those who work at their own businesses in the total number of all workers).

The question about whether business was new or not was not asked in the 1995, 2006, and 2008 surveys. Most of our analyses therefore start with the year 1997. The identification of founders and workers of startups/established firms in 2006 and 2008 (used for showing illustrative graphs only) is based on a well-established imputation method (see, e.g., Little and Rubin, 2019). The details of the imputation method are in Appendix. We deliberately focus on the top tail of the distribution of scientists and engineers and their decisions on whether to found a startup. The sample therefore excludes other types of science-based firms, for example those owned and operated by individuals with Bachelor's or Master's degrees in science and engineering, or those owned by individuals without degrees in science and engineering.

Self employment is often associated with motivations other than launching high-opportunity startups (e.g., Hurst and Pugsley, 2011). Business incorporation has lately become a popular indicator of the growth potential of startups (see e.g. Åstebro and Tåg, 2017; Levine and Rubinstein, 2017). Therefore, in our main analysis we focus on owners of *incorporated* startups, whom we call "founders." This term refers to owners at the time observed in the data, whose incorporated startups were their full-time primary employer and who drew non-zero salaries. We also sometimes use the term "startup rate" as a shortcut when referring to the share of founders among all incorporated business owners.

To track the evolution of the number of R&D and managerial job tasks we use SDR data that contain responses to the question "Which of the following work activities occupied at least 10 percent of your time during a typical week on the job?" The list of possible alternatives includes 14 different activities. See Table A2 (all figures and tables starting with the letter "A" are in the Appendix). We call these "tasks" and follow the classification suggested by the SDR to identify four of those activities as "R&D" tasks: "basic research," "applied research," "development," and "design." We add up tasks in the R&D category to obtain the number of R&D tasks performed by founders and workers. We also follow the SDR-suggested classification to classify five work activities as management and administration (abbreviated as "management" tasks) as detailed in Appendix.

As in recent empirical work (e.g., Tåg, Åstebro & Thompson, 2016), We infer firms' hierarchical structure based on microdata. SDR provides data on the number of people supervised directly and indirectly for those that respond to the survey. Consistent with prior work we call the former "span of control" and the latter "depth of the hierarchy" as it relates to the respondents' job position (Guadalupe, Li & Wulf, 2014).

The Dynamics of Science-based Startups

Figure 1 displays our key motivating findings for this study: a large decline over the past two decades in 1) the share of startup founders among all business owners, and 2) the share of workers employed at startups. Together, these imply a large decline in startup rates within the most science-intensive part of the workforce in the U.S. economy. Specifically, in 1997, 34 percent of business owners reported that their businesses had been founded within the previous five years, but by 2017 this rate had declined to 21 percent, a decline of 38 percent. Meanwhile, the employment share of science and engineering PhDs at startups as a fraction of all employment opportunities in private firms had fallen from 12 percent in 1997 to 8 percent in 2017, despite two temporary employment spikes in 2001 and 2008.

--- Figure 1 around here ---

The afore-mentioned spikes could potentially be related to abnormal events in computers and information technologies. But as computer-related occupations represent a relatively small minority of PhDs in our sample (Table A5), their impact on the overall trends is in fact marginal. We also reconstructed Figure 1 excluding computer-related occupations and found the same trends (Figure A9). It is still possible that the dot-com boom incentivized many recent PhDs with coding skills to work in startups, while the 2008 financial crisis may have temporarily pushed many recent graduates out of the banking and finance sectors and towards startups. Whatever the explanation, these temporary shocks appear not to have affected the overall declining trends in PhDs deciding to start or work for new firms.

Importantly, the downward trend is universal and found across all categories of PhDs in our data. The declines in the fractions of founders from 1997-2017 for some subcategories are: 39 percent for males, 36 percent for Whites, 35 percent for those residing in California, 41 percent for those whose work activity included computer-related tasks, and so on (Figures A2-A8). One category of founders that suffered an especially large decline is the non-US born PhD degree holders where the fraction of founders was higher than in the whole sample in 1997 (at 41.5 percent), but had fallen to about 21 percent by 2017 (a decline of almost 50 percent). This exaggerated drop may be associated with a tightening of US visa requirements (Roach and Skrentny, 2019) but notably, it is still in line with the overall trend. We also examined the evolution over time of the composition of founders with PhDs in terms of demographics, fields in which they majored, and so on, but found nothing in these compositional data that could explain the sharp decline in the share of PhDs that become founders of startups (Figures A10 and A11). The fraction of PhDs who decide to work in academia is also very stable over time and cannot therefore explain the steep drop in founders and employees with PhDs in startups.

Life sciences-related startups have perhaps undergone the most dramatic changes over the last 20 years so we thought that they may perhaps buck the overall trend. Slightly over 20 percent of founders in our data have PhD degrees in life sciences (including bioengineering and biomedical engineering). It turns out that the trend toward declining startup founding rate is similar and perhaps even more pronounced for startups with PhD founders in the life sciences. Specifically, the fraction of founders among PhD degree recipients in life sciences went down by almost 50 percent from 1997-2017 (Figure A12).

A Potential Explanation for the Decline in Science-based Startups

Preliminary Evidence

What is behind this decline in science-based startups? Employing individual-level work data provides an opportunity to examine one explanation that has not yet attracted much attention in the literature. As mentioned in the introduction, there has been an increasing complexity in conducting science-based innovation (a.k.a. the "burden of knowledge"), documented in Jones (2009) and more recently in Bloom, Jones, Van Reenen & Webb (2020). We focus on the potential implications this may have for the attractiveness of founding a startup as opposed to working at an established firm. The underlying idea is that an increasing scientific complexity could be especially hard to handle by founders in science-based startups. In what follows we tease out some more specific implications that can then be taken to our data.

We first note an increase in post-PhD experience for science-based startup founders, which goes up on average by more than a full year (from 16.4 to 17.5 years) from 1997-2017 in the raw data. Regression estimations controlling for gender, ethnicity, occupation and the U.S. state of employment imply that the average founder had about 15 percent longer post-PhD experience in 2017 than in 1997, statistically highly significant (Table A6).

Figure 2 shows another closely related fact—there is a pronounced decline over time in the earnings of less experienced founders relative to their more experienced peers. Here, we separate founders by the median number of years after PhD (13 years), although other reasonable cut-offs lead to similar results. Between 1997 and 2001, founders with experience at or below the median earned more than the rest of the sample (consistent with Ouimet and Zarutskie, 2014), but by the 2010s, the situation is reversed, with less experienced founders earning on average 30-40 percent less than experienced founders. The earnings of less experienced founders declined also in absolute terms. The average inflation-adjusted earnings of founders with below the median post-PhD experience were $72,616 in 1997, whereas 20 years later their earnings were $57,517, a decline of more than 20 percent (Table A7).

--- Figure 2 around here ---

When viewed together, these two findings do not seem to be consistent with the simple "aging" hypothesis of the potential startup population. If, for example, the increase in the post-PhD experience of startup founders was mainly driven by the dwindling supply of younger potential founders, there would be no reason for the earnings of younger founders that do appear in the data to go down relative to the earnings of more experienced founders, let alone for an absolute decline in such earnings noted above. Instead, the evidence is consistent with the presence of factors operating on the "demand" side, namely, those coming from increasing knowledge complexity demands. As noted by Jones (2009), more work experience is akin to prolonging the training in preparation for shouldering the burden of knowledge (see also Braguinsky, Klepper & Ohyama, 2012). Increasing work experience among science-based founders over time would be exactly what we would expect to see in the data if the burden of knowledge were a major factor driving the decision to found a startup. Moreover, in this case, it is also easy to understand why experienced founders would develop an increasing earnings advantage over less experienced founders over time.

The SDR data offer a unique opportunity to probe the hypothesis further by looking at trends in the nature of work. Specifically, we examine the number of different tasks performed, especially R&D tasks that founders have to deal with.

Figure 3, Panel A, illustrates the dynamics of the number of tasks reported by founders. The total number (with a maximum of 14) is measured on the left-hand scale whereas the numbers of management and R&D tasks, whose maxima are five and four, respectively, are measured on the right-hand scale. The average number of all tasks for founders increased by about 15 percent from the beginning to the end of our sample, a statistically significant difference using a double-sided *t*-test. Furthermore, although the number of both R&D and management tasks increased, the increase is more pronounced in R&D, for which it rose by more than 50 percent from 1997 to 2017. The increase in R&D tasks was equally pronounced at all experience levels (column 1 in Table 1 below). As a result, R&D tasks that comprised about 25 percent of all tasks conducted by founders in 1997 increased to 34 percent of all tasks in 2017. This was not accompanied by any decline in management tasks, so the founders had to shoulder the burden of doing more R&D tasks while also performing the same or more management tasks.

--- Figure 3 around here ---

Panel B of Figure 3 contrasts the above findings with the trends for workers at established firms. The number of R&D tasks increase by about 12 percent from 1997 to 2017. The number of management tasks handled by workers, however, remained relatively flat. Also, workers had to deal with significantly less management tasks than founders in terms of levels: about 25 percent less at the beginning of the study period, with the gap increasing towards the end of it.

Compared with an established firm, managing a startup is likely to constrain founders' ability to organize the firm's knowledge hierarchy efficiently at least until it has succeeded in growing well beyond its initial size and financial constraints have been reduced. Lacking an efficient allocation of skills in the hierarchy, this constraint implies that founders would take on more work tasks themselves when the complexity of innovation increases, and this is consistent with the evidence in Figure 3.

Formal Analyses

The evidence so far suggests that science-based startups, especially those founded by less experienced PhDs may have struggled to effectively manage the increasing complexity of innovation. To examine this more rigorously, while accounting for various other differences between founders and workers, we turn to regression analysis. The estimation equation is:

$$y_{it} = \alpha + \beta_1 t + \beta_2 exp_{it} + \gamma \cdot X_{it} + \varepsilon_{it}, \qquad (1)$$

where y_{it} denotes outcome variables for individual *i* at time *t*, which are (a) the number of R&D tasks, (b) the span of control, and (c) the depth of the hierarchy; t is the time trend; exp_{it} denotes experience (years after PhD) of individual *i* at time *t*; $\boldsymbol{X_{it}}$ is a vector of controls, consisting of demographic, occupational, and geographic location dummies; and ε_{it} is the error term. We employ pooled cross-section OLS because startups are new businesses that are "redrawn" in each survey, and we estimate equation

(1) separately on the samples of founders and workers at established firms.[4] Estimation results are presented in Table 1.

--- Table 1 around here ---

The coefficients on the time trend variable in columns (1) and (2) in Panel A confirm that the number of R&D tasks for both founders and workers at established firms is trending up, while trending stronger for founders (the coefficients for founders and workers are highly statistically significantly different). Thus, one might conclude that the increasing burden of knowledge to some extent affects all science-based workers; both founders and employees at established firms. However, coefficients on work experience show that workers at established firms, but not founders, tend to perform fewer R&D tasks as their experience grows (the difference is highly statistically significant). Thus, established firms but not startups appear to employ a greater division of labor in innovation by assigning more experienced workers to team leadership roles. In Panel B I dig deeper into this by looking at the span of control and the depth of the hierarchy. Neither of these changes over time for founders. Most startups are small, so opportunities to restructure to counter the increasing complexity in innovation must be inherently limited.

In contrast, both the span of control and the depth of the hierarchy (especially the latter) strongly increase with experience for workers with science backgrounds at established firms, in line with such workers performing fewer R&D tasks as they age. Significantly, the negative coefficient on the time trend for the span of control in

[4] In recognition of inherent difficulties with cross-sample comparisons of entrepreneurs and workers, especially in survey data (see, e.g., Hamilton, 2000), all the analyses here and below are based on comparing time trends *within* each sample of founders and workers, respectively.

column (4) indicates that the span of control *decreases* over time for workers at established firms, on average by about 0.46 percent per year (statistically highly significant). In contrast, the coefficient on the time trend for the depth of the hierarchy in column (6) is positive and statistically also highly significant, indicating that hierarchies *deepen* over time for workers at established firms on average by about 0.32 percent per year. Together, the results suggest that established firms cope with the increasing complexity of innovation by introducing additional layers of hierarchy and by narrowing their science-based workers' scope of work.[5]

We probe this line of reasoning further by looking at large firms. If the increasing complexity of innovation is indeed a culprit but can be mitigated by introducing additional layers of management, large firms should have greater opportunities to implement these organizational changes because they can more easily absorb the additional fixed cost of additional management layers (Caliendo and Rossi-Hansberg, 2012). Hence, we re-estimate equation (1) for workers at large firms. SDR surveys contain responses to the question about broadly defined size categories of employers, measured by the responders' estimates of how many people worked for the principal employer across all locations. There are eight categories, ranging from 10 or fewer employees to 25,000+ employees. For the purpose of this analysis, we define firms in the two largest categories, 5,000-24,999 employees and 25,000+ employees,

[5] Suppose that a firm with two R&D teams with 10 members each added a second layer made up of a single manager, promoted from a supervisory position on one of the teams. The newly promoted manager had nine direct reports before the change but now has zero direct and 19 indirect reports. Her replacement will have eight, not nine, direct reports. More generally, reorganizing a firm by adding layers of hierarchy while holding the firm size constant can be seen as accompanied by a decrease in the number of direct reports and an increase in the number of indirect reports.

as “large firms.” This category accounts for about half the workers at established firms in our sample.

In Table 2, columns (1), (3), and (5), we present estimation results from eight surveys in which the question about employer age was asked, so the sample is limited to employees at large firms that we know are not startups. In columns (2), (4), and (6), we employ data from all 11 available surveys under the assumption that firms with 5,000 or more employees are highly unlikely to be startups. Results from all surveys and from the eight surveys where we know firm age are very similar.

--- Table 2 around here ---

Comparing the estimation results in Table 2 to those in Table 1, we should notice that the coefficient on the time trend variable in the number of R&D tasks is smaller among workers at large firms compared with all workers and is about half the magnitude of the corresponding coefficient for founders in column (1) of Table 1 (all difference mentioned are highly statistically significant). Thus, workers at large firms faced an increase in the burden of knowledge at about half the pace of founders. Perhaps even more interestingly, the coefficient on the time trend variable in the regression where the dependent variable is the depth of the hierarchy is twice the magnitude of the coefficient in the same regression for the sample of all workers at established firms in Table 1. Thus, to offset the increasing burden of knowledge, large firms create deeper hierarchies at twice the rate compared to all firms.[6]

[6] Recall that our analysis concerns workers with PhDs, and so primarily these results refer to changes in the hierarchy for those working on scientific tasks.

As larger firms find better ways to cope with the burden of knowledge, PhD degree holders should naturally start choosing to be employed by such firms more. The flip side of this would be a decrease in the share of PhD degree holders employed by small firms. Elfenbein et al. (2010) find that experience working at small firms is closely related to the propensity of subsequently founding startups; therefore, decreasing share of employment in such firms could indirectly contribute to decreasing science-based startup rates.

To see how important this factor may be quantitatively, in Figure 4 we present the dynamics of the shares of employment of PhDs in science and engineering at large (5,000 or more employees) and small (less than 25 employees) for-profit firms, using all available surveys (here, firm age is not an issue). The share of PhDs employed at small firms has dropped by about a quarter (from around 12 to around 9 percent) since the late 1990s. The share employed in the smallest firms (less than 10 employees, not shown) has seen similar declines. The trend for PhDs to choose employment in large firms as opposed to small firms may therefore be contributing to lower startup rates not just directly but also indirectly by narrowing opportunities to "learn" entrepreneurship through small-firm employment.

--- Figure 4 around here ---

We also confirmed the evidence on founders' earnings in regression analysis which includes various controls. In Table 3 the estimation equation is similar to Equation (1) except that the outcome variable is the (logged) salary of individual *i* at time *t*, while the explanatory variables of interest include the (IHS-transformed) number of R&D tasks, the span of control, and the depth of the hierarchy, and there is an interaction term between the time trend and experience.

Founders' earnings decline on average by about 1.6 percent per year (column 1, row 1, statistically highly significant). However, as we already saw, this is offset by an opposite trend in returns to experience. The mean number of years after PhD in the sample of founders is 15.9 years; hence, the coefficient on the interaction term between experience and time trend implies that the negative baseline time trend is offset at the mean work experience. In contrast, workers' real earnings grow over time regardless of experience (column 2, row 1, statistically highly significant).

--- Table 3 around here ---

Workers at established firms are also rewarded for taking on more tasks. Doubling the number of R&D tasks increases earnings for workers at established firms by 4.4 percent. The point estimate for founders is similar, but the standard errors are high so the most I can say in this case is that, controlling for all other factors, there is large variance in outcomes.

Evidence from R&D-focused Founders

If increasing complexity of innovation is indeed a major factor behind the decline in science-based startups, we would expect the patterns documented above to be magnified among founders whose startups are the most R&D-intensive. We use this logic to evaluate our hypothesis further using some additional data from the SDR surveys.

In addition to the list of work activities, SDR asks the question “On which work activity did you spend most of your time during a typical week on the job?” We call founders who listed R&D as their primary work activity “R&D-focused” founders and similar for owners of established businesses and workers at established firms.

To begin with, Table 4 presents the dynamics of the share of R&D-focused founders, R&D-focused owners of established incorporated businesses, and R&D-focused workers at established firms in the total number of individuals in each category, as well as the share of R&D-focused workers at all large firms using all available surveys (regardless of business age). The share of R&D-focused founders among all founders in column (1) increases over time—from 26 percent in 1997 to 32 percent in 2017, an increase of 21 percent. However, the share of R&D-focused business owners among owners of established businesses in column (2) rises even faster, from about 12 percent in 1997 to over 30 percent in 2017, suggesting that the burden of knowledge is widespread. Thus, when we split the data between startups and established own businesses as before, the rate of R&D-focused startups steeply declines, and this decline, from 54 percent in 1997 to 20 percent in 2017, is larger in magnitude than the decline in the overall startup rate shown in Figure 1.

--- Table 4 around here ---

Note that an R&D-focused founder (owner) and a startup (own established business) engaged in R&D might not be the same thing. A business might be primarily engaged in R&D whereas the owner primarily manages a team of researchers. An increasing burden of knowledge, leading to an increasing number of task execution exceptions brought to the attention of the owner might change this, shifting the primary work activity from management to R&D. In fact, this is the interpretation that appears to us to be most consistent with the large increase in the share of R&D-focused owners among established incorporated own businesses. That is, the fact that owners of so many established businesses tend to report R&D as their main work activity later in the sample could be driven by an increasing number of exceptions brought to the attention of the owner rather than a reprofiling of the business itself. Note also that, as can be seen in columns (3) and (4) of Table 4, the share of R&D-focused workers at established firms remains stable and even appears to decline slightly in the sample of workers at large firms. This further favors interpreting the increase in the fraction of R&D-focused business owners as evidence of the need for them to deal with an increasing number of exceptions because of the rising complexity of innovation and lack of recourse to taller hierarchies.

If the increasing burden of knowledge is indeed an important driver of these trends, it should also be reflected in steeper earnings decline for especially less experienced R&D-focused founders compared to other founders. Table 5 presents the estimation results from the same regression for which results were in Table 3, but with the sample limited to R&D-focused founders. The negative earnings time trend is indeed much more pronounced in Table 5 than it is in Table 3—the decline at the baseline is more than 4.4 percent per year and statistically highly significant. Notably, comparing the coefficients on the interaction term between experience and time trend in Tables 3 and 5, the returns to experience are increasing over time for R&D-focused founders at twice the rate of all founders while the point estimates of the returns to the span of control also indicate that R&D-focused founders can increase their earnings by much more than other founders if they can lead larger teams. Thus, more experienced R&D-focused founders as well as those who manage to quickly overcome the initial founding team size constraint appear to be in a good position to compensate for the increasing burden of knowledge that haunts their less experienced and smaller-sized peers. We will come back to this issue in the next section where we examine the trends in the quality of human capital of science-based startups.

In summary, we find that while R&D-focused owners are becoming more plentiful among both science-based startup founders and established business owners, this is likely to be driven by such science-based business owners having to shoulder more R&D tasks on their own due to the relative lack of opportunity to delegate R&D tasks in startups. This explanation appears to be most consistent with the fact that both the startup rates and earnings' declines are even more pronounced among R&D-focused founders than among all founders.

--- Table 5 around here ---

Who Chooses Startups?

The Human Capital of Founders and Startup Quality

Some research has found a declining trend in cognitive skills among new entrants (Ayyagari and Maksimovic, 2017). Our previous analysis suggesting increasing burden of knowledge as a reason for the decline in science-based startup ratios appears to lead us to a different conclusion—if knowledge-related demands on founders are growing, the human capital of those who still choose to found them should be going up, not down. We already saw this to be true in the previous section in terms of pre-founding work experience. Here we probe this issue further by examining the changes in the human capital of founders over the time span of our analysis.

SDR surveys have a (partial) longitudinal aspect, as each new survey between 1995 and 2013 returned to previous survey respondents, while new PhD recipients were added to every survey. The attrition rate between adjacent surveys is about 10-15 percent. In 2015, however, the survey was sent to an almost entirely new sample of respondents (the attrition rate from the 2013 to the 2015 survey is about 90 percent). Hence, we use two waves of panel data available for founders: the 1997-2013 wave and the 2015-2017 wave.

To estimate changes in the human capital of founders we use the longitudinal aspect of the data and look at all those employed at a private-for-profit business in a survey. We create a dummy that equals one if the individual is reported to be a startup founder in the next survey and zero if he or she remained employed at a private-sector for-profit business in the next survey as well. We then use a standard approach (Choi, Goldschlag, Haltiwanger & Kim., 2021; Azoulay, Jones, Kim & Miranda, 2020) with prior ("pre-choice") salaries as a measure of the human capital of future founders to obtain an estimate of the human capital differential between future founders and those who remained in paid work for other firms, based on how much they earned while both types of individuals were employed for pay. Because the demographics are different across future founders and those who remained in paid work for other firms (in particular, the former tend to have more experience than the latter), we employ predicted logged pre-choice salaries from an earnings regression estimated separately on the data from each relevant pre-choice survey (1995, 1997, 1999, 2001, 2008, 2010, and 2015), in which the dependent variable is the pre-choice logged salary, and the explanatory variables are work experience and its squared term, tenure on the current (pre-choice) job, gender and ethnicity dummies, as well as 75 occupation dummies, U.S. state of employment (Massachusetts, New York, California, etc.) dummies, and eight employer size-category dummies. Figure 5 plots these predicted logged pre-choice salaries and the 95 percent confidence intervals.

--- Figure 5 around here ---

As seen in Figure 5, the average predicted earnings of those who become founders in the next survey are lower than the average predicted earnings of those who remain at private sector for-profit businesses in the 1990s. However, the predicted earnings converge in 2001 and then the average predicted earnings of those who become founders exceed those of workers who continue to work for other firms in 2008-2010. Because of the small sample size of founders in these longitudinal data, the 95 percent confidence intervals for the most part overlap. The only time when there is a statistically significant difference at the 5 percent level between predicted pre-choice salaries of these two categories of individuals, in 2010, the pre-choice predicted salaries of future founders are actually higher than those who continue to work for other firms. This suggests that PhDs who left paid employment to start their own incorporated businesses between 2010 and 2013, that is, right after the Great Recession, had significantly higher human capital than those who remained in paid work.

In Figure 6 we look at an alternative measure of human capital—the share of individuals who graduated from top-ranked PhD programs. SDR data contain Integrated Postsecondary Education codes for the institutions from which the respondents graduated and their field of specialization in the graduate program. I matched these codes with the National Research Council (NRC) PhD program rankings published in 1993 and constructed a dummy that equals one if the individual graduated from a PhD-granting institution in the top 15 in the NRC rankings and zero otherwise. We also added to these top-ranked programs graduates whose specific PhD area was among the top 15 in the NRC's area rankings (the areas were mathematics and physics, engineering, social sciences, and biological sciences). Figure 6 is consistent with Figure 5 in showing that founders lag behind those who continue to work for pay in terms of the share graduating from top-ranked programs in the 1990s but catch up to a large degree after 2003. Notably, the catchup trend is even more pronounced among founders with below the median years after PhD, despite the evidence presented that they suffer an increasing earnings penalty over time.

--- Figure 6 around here ---

Our findings indicating increasing human capital of founders over time is consistent with a "selection effect" due to higher barriers to entry that, according to our story, stem from the increasing burden of knowledge. Note that this selection effect does not have to lead to an increase in average earnings of the remaining founders, because the increasing burden of knowledge reduces earnings at all levels of human capital.

The increase in human capital of founders in recent years translates into better startup growth, however. As mentioned, most startups start small. Despite higher human capital of founders in more recent years, there is not much of a trend toward higher initial startup size—the fraction of businesses with fewer than 10 employees among all startups was 85 percent in 1997 while that fraction declined to 82 percent in 2017, with some fluctuations in-between. However, those that survive tend to grow faster in more recent years.

To examine growth of surviving startups, we construct panel data of startup founders. Specifically, we track founders, retaining those who manage their own incorporated businesses between two adjacent surveys. Figure 7 shows changes in the average size of startups from one survey year to the next.[7] Firm size is classified into eight categories, with category 1 representing less than 10 employees. The figure displays that surviving startups are larger in the period 2003-2015 than they are in the period 1997-2001. Moreover, growth rates accelerate over time: growth is between 3-5 percent for startups first observed between 1997 and 2001 but then increase and reach 10 percent for startups first observed in 2015. The increase in growth rates for surviving startups is confirmed in a regression using repeated observations on founders across two adjacent surveys. The coefficient on time trend implies that growth rates are increasing by about 1.6 percent per year, after controlling for the initial startup size and individual fixed effects, statistically significant at the five percent level (Table 6). Thus, there might be a silver lining: while increasing complexity of innovation is associated with a reduction in the number of startups, those that do form are started by individuals with higher human capital, they tend to be larger in size initially, and they also tend to grow faster over time.

--- Table 6 around here ---

[7] Since we need repeated observations from two surveys, we had to drop 2017 which is the last year in the data.

Changes in Job Motivation and Job Satisfaction

Some SDR surveys also contain questions about the importance of various factors when thinking about a job, with the respondents asked to check one of the following boxes for each factor: “very important,” “somewhat important,” “somewhat unimportant,” and “not important at all.” See Appendix for the list of all factors in the surveys.

Figure 8 shows a key insight that can be gained from these answers. In this figure, we focus on the fraction of those who responded that salary was very important versus the fraction of those who responded that the degree of independence was very important and compare them across startup founders and workers in established firms, and across time.

Not surprisingly perhaps, the degree of independence is listed as very important relative to salary at much higher rates among startup founders than among workers (Hamilton, 2000; Roach and Sauermann, 2015; see Table A8 for absolute numbers in each category). More interesting is the dynamics of these numbers. For workers at established firms, the relative importance of the salary factor steadily grows over time relative to the degree of independence and it exceeds the importance of the degree of independence after 2013 for all workers and already after 2010 for young workers (those below the median number of years after PhD). Significantly, young workers list salary as a relatively more important factor compared to independence than do all workers.

The trends among startup founders are even more interesting. For all startup founders there is either no trend or perhaps a slight trend towards increasing relative importance of the salary factor, similar to workers in established firms. Among young startup founders, however, there is an opposite trend where the relative importance of the degree of independence becomes stronger compared to the salary factor over time. The difference between the fraction of those who listed independence as very important and those who listed salary as very important increases on average by approximately 50 percent from the start to the end of our sample.

Another interesting insight can be obtained from the answers to questions about satisfaction with various aspects of the respondents' principal job. The list of aspects in those questions mirrors the list of factors that are important when thinking about a job and, once again, the respondents are asked to choose between "very satisfied," "somewhat satisfied," "somewhat dissatisfied," and "very dissatisfied" (see Appendix). As can be seen from Table A9, there is a steep decline in the fraction of those satisfied with their salary on the principal job among young startup founders. This is consistent with the trend in their real earnings shown in Table A7 (it also provides a cross-validation of both findings). However, the satisfaction with the degree of independence is relatively stable at a high level, and even after a steep decline in the 2010s, the degree of satisfaction among young startup founders with their salary on the principal job is not that much below the degree of satisfaction among young paid workers with their salary in established firms.

In sum, we may conclude that new demands presented by the burden of knowledge have raised the requirements for founders' human capital and that market selection increasingly "chooses" PhD degree holders who place especially high value on independence to launch science-based startups. In contrast, those who choose to work for established firms display an increasing preference for salary.

Robustness and Alternative Explanations

To gain insights about the dynamics of startups most likely to contribute to innovation and growth in the economy, we studied founders with a PhD in science and engineering who started an incorporated business. In this section, we briefly examine what happens if we perturb that measure and also consider some more potential alternative interpretations of our findings. Note that in previous sections it has already been showed that our findings hold across all demographics, PhD fields and location choices as well as that they are robust to excluding computer and IT-related fields and occupations, which represent a small minority of PhDs in science and engineering anyway.

Type of Entrepreneurial Opportunity

Following a series of articles showing that incorporated businesses pursue the highest opportunities, we chose to analyze founders of incorporated startups. However, including founders of unincorporated startups in the analysis yields similar findings, provided the sample is limited to founders whose job and education are closely related (Braguinsky et al., 2012; Ohyama, 2015). The details are in Appendix.

Declining supply of PhDs?

Could the decline in the creation of science-based startups somehow be driven by a decline in the number of PhDs? It appears not. The overall supply of recent science and engineering PhD recipients is actually going up, as the number of PhDs awarded in science and engineering fields increased from 28,486 in 1997 to 41,294 in 2017 (National Science Foundation, 2019).

Aging of the population?

Over the past few decades, the U.S. population has been aging. Since young people tend to disproportionately found or join new businesses (Ouimet and Zarutskie, 2014), the aging of those receiving PhDs could explain the decline in Figure 1 (Hopenhayn et al., 2018; Karahan et al., 2019). However, if we define "young" as those below the median age of 45 (or any other reasonable cutoff, for that matter), the share of such young founders in the total number of all founders was stable, about 35-40 percent throughout the period studied. Furthermore, while the age at which individuals received their PhDs in science and engineering had increased steadily until about 1993––the average recipient was 28.7 years old in 1971 and 31.3 years old in 1995—this trend was halted and even reversed after that—the average recipient in 2015 was 30.3 years old (Figure A15). Apparently, the age of founders is not strongly associated with the decline in science-based startups.[8]

Declining Venture Capital?

A simple explanation for the decline in science-based startups is that the aggregate amount of venture capital for startups has declined. However, the supply of venture funding has instead substantially increased since the dot com crash. For example, the amount of global venture capital investments increased tenfold, from $25 billion in 2004 to $254 billion in 2018 (National Venture Capital Association, 2019).

[8] In contrast, as mentioned, accumulated work experience (years after PhD) prior to founding a startup did tend to increase over time. See detailed discussion in Appendix.

More than 8,380 venture-backed companies received $131 billion in funding in 2018, surpassing the $100 billion mark set at the height of the dot-com boom in 2000. For more detail on how venture capital has shifted over the years, see Lerner and Nanda (2020). Nevertheless, the recent decline in VC disbursements may potentially accelerate the trends observed.

Discussion

Human capital is argued to be a key source of firm competitive advantage (Castanias and Helfat, 1991, 2001). But the same attributes that make employees strategic assets also create strategic dilemmas (Coff, 1997). We show in this article that startups face increasing pressure to motivate PhD-educated founders and their workers to create, join, and stay in startups. Startups appear to have been in a losing battle. To this effect, we first establish a new fact. Since 1997, the share of startup founders with PhD degrees in science and engineering among all independent incorporated businesses owned by individuals with PhD degrees in science and engineering in the United States declined by around 38 percent, not limited to any particular founder demographic, ethnic group, region, occupation, or field of science. The declines are even larger among foreign born and minority founders. The employment share at startups has followed the same path of decline.[9] The trends appear despite a massive increase in

[9] As mentioned, we follow common practice (e.g., Haltiwanger, Hathway, and Miranda, 2014: Decker, Haltiwanger, Jarmin, and Miranda, 2014) and define startup rates (and employment in them) as the share of new incorporated businesses with PhD founders in all incorporated businesses with PhD owners (and employment in them). The point is to examine the trend in where Ph.D.s decide to apply their skills; startups or established firms. We do not study the share of self-employment among all workers, and we do not study the absolute number of startups.

venture capital funding and an impressive absolute growth in the number of PhDs graduated. There are similar trends in startups in the high-tech sector and among inventor-founders (Goldschlag and Akcigit, 2022; Goldschlag and Miranda, 2020). The trending decline is large enough to contemplate if there is a need for managerial (or public policy) action. In order to do so one first need an answer to the question “Why is this happening?”

To address this question, we set out on a journey of abductive reasoning (Pillai et al., 2021; Shani et al., 2020; Tohmé and Crespo, 2013), to produce exploratory hypotheses and yield plausible explanations about these puzzling phenomena. Using individual-level work data we show that founders in science-based startups have had to perform an increasing number of R&D tasks over time while also increasing the number of management tasks that they must perform. Nevertheless, their earnings generally have not reflected the added workload, as average earnings have been declining, although more experienced founders have seen an increase in earnings.

We also documented that their outside option to go and work for established firms has instead considerably improved along the same dimensions that have deteriorated in startups. Established firms, and especially larger ones, have had more room to adjust their innovative human capital management practices, in particular by introducing more layers in their hierarchies, reducing knowledge workers’ span of control, and allocating more experienced workers to positions with greater managerial responsibility. Hence, they attract recent PhDs in greater numbers and therefore, less of these tend to create and join startups.

The totality of the evidence uncovered in this article points to an increasing complexity of innovation as a plausible explanation for the decline in science-based startups. The natural limits imposed by running a small firm with little opportunity to delegate R&D work and an already high and increasing degree of multitasking by the founder has made it more difficult to manage innovation in science-based startups. More experienced founders are therefore sought, and they fare better in terms of earnings. However, as the supply of new founders appears dwindling, the pool of experienced founders may also eventually dry up too.

Why should one care that innovative work has gotten, relatively, more comfortable in large firms? While a secular decline of new firms in mature industries might be a normal function of industry evolution (Klepper, 1996; Agarwal et al., 2020), such a decline in new, high-tech industries is a relatively recent phenomenon (Haltiwanger, Hathaway & Miranda, 2014). This can have profound implications for innovation. After all, it is new firms from which many new technologies and business opportunities typically originate (Schumpeter, 1942) so that "[a] country cannot be great over a sustained period without a steady flow of great new firms" (Klepper, 2016, p. 62).

Our findings do suggest that there might be a silver lining: science-based startups that are still being founded appear to be larger and grow faster, and their founders tend to have not just more experience but also relatively higher human capital over time. The data also suggest that the recent declines in PhD founders are jointly driven by changes in pecuniary compensations and changing job tasks. It is plausible that the decline in the fraction of PhD founders and employees may flatten out as there are strong intrinsic motivations both to becoming a scientist and for a scientist to become a founder (or joiner) of a startup (Stern, 2004; Roach and Sauermann, 2015). Recall that there is not just a trend toward higher human capital of the remaining founders, but also a tendency to be relatively more motivated by non-pecuniary motivations in more recent years. Those motivations would continue to fuel the creation of startups in the future.

That said, if the noted increase in pay difference remains, there could still be a permanently lower supply of willing PhD founders in the future. If the goal is to restore business dynamism in the high-tech sector, alleviating the negative impacts of the increasing complexity of innovation should perhaps be front and center in the strategy to attain it. Managerial practitioners and policy makers may need to find ways to make the job of the founder more attractive to keep the flow of commercialization of new ideas coming. For founders, that may mean more training in multitasking and use of advanced management practices and technical management tools. Another alternative is more, and better lubrication of the market for good and experienced "natural" founders, and there are some indications in the data that this is happening. The findings in this article point to the importance of secularly increasing value of complementary assets in work practices, especially for founders and employees that are high value generators. An alternative view is that the increasing dominance of a few very large tech firms that the increasing burden of knowledge is fueling is just fine, and that startups need not play an important role for economic development.

One question remains as an avenue for future research is whether the increasing supply of PhDs would also reduce the fraction entrepreneurs, including startup founders and owners of established businesses, among all doctorate recipients in science and engineering. While we are focusing on the startup rate among PhDs in this chapter and thus have nothing to say about the fraction of entrepreneurs among all PhDs, this is an important question to be addressed in its own right by future research.

The data we employ in this study is individual- and work-based. In particular, there is very little firm-level data, which limits our ability to test alternative hypotheses. Note, however, that the level of detail about work activities available in the SDR surveys is not available in the Census (LEHD or even ACS) data and there is no known bridge between SDR surveys and Census data. Thus, as it is, the conclusions from individual-level work data at the very least present an important novel explanation to consider for the declining dynamism in science-based startups. Future research might be able to overcome current data limitations and bring together firm-level and individual-level analyses for an even more comprehensive picture of what is going on in the most technologically important parts among U.S. businesses. In the next chapter of this dissertation, I take steps toward that goal by employing the U.S. Census data.

Chapter 2: Complex Innovation and the "Visible Hand": the Role of Knowledge Interdependence in Employee Entrepreneurship[10]

Introduction

As mentioned in the previous chapter, the U.S. has experienced a generally declining trend in entrepreneurship and business dynamism, as manifested in declining firm entry rates, falling employment share in young firms, and declining job mobility, especially from established to new firms (Decker et al., 2014; Akcigit and Ates, 2019). However, whether such a trend is an ominous concern depends on its underlying causes. It would be a grave concern if it is primarily driven by the increasing market power of large corporations not necessarily rooted in innovation but stifling the entry of young firms from which most new technologies and business opportunities typically originate (Schumpeter, 1942). However, it might be less of a concern if this trend is driven by technological changes that simply require an efficient (re)allocation of resources from startups to large incumbent firms. That is to say, the nature of technological change may be transforming the roles of large established firms to become relatively more important hotbeds for innovation and the engine of economic growth as opposed to startups.

[10] This research uses data from the Census Bureau's Longitudinal Employer-Household Dynamics Program, which was partially supported by the following National Science Foundation Grants SES-9978093, SES-0339191 and ITR-0427889; National Institute on Aging Grant AG018854; and grants from the Alfred P. Sloan Foundation. Any views expressed are those of the author and not those of the U.S. Census Bureau. The Census Bureau's Disclosure Review Board and Disclosure Avoidance Officers have reviewed this information product for unauthorized disclosure of confidential information and have approved the disclosure avoidance practices applied to this release. This research was performed at a Federal Statistical Research Data Center under FSRDC Project Number 2296. (CBDRB-FY22-P2296-R9700)

This dilemma may not be that new as it is clearly present already when comparing the writings of early and later Schumpeter (Schumpeter, 1911, 1942). The intellectual debate about the roles of Adam Smith's Invisible Hand and Alfred Chandler's Visible Hand in innovation and technological progress, in particular, is indeed quite old. However, it may be worthwhile revisiting it once again, as scholars have noted a prominent change in scientific development and innovation in recent decades. Specifically, there is an increasing trend in the burden and complexity of knowledge necessary for innovation (Jones, 2009), and new ideas are getting "harder to find" (Bloom et al., 2020). However, there appears to have so far been little attempt to examine a possible connection between the increasing complexity of knowledge and the aforementioned trends toward reduced entrepreneurship and business dynamism. It is tempting to conjecture that more complex innovations could be more efficiently handled by more complex (larger) organizations, in which case small startups may be facing an increasing relative disadvantage in conducting complex innovation compared to large, established firms. Indeed, in the first Chapter of this dissertation, my coauthors and I found evidence to this effect using detailed work data from the NSF Surveys of Doctorate Recipients but the data we used there do not allow us to see what is going on inside the firms and, in particular, where startup founders come from.

To fill this gap, I leverage matched employee-employer data from the U.S. Census Bureau and adopt an abductive approach (Heckman and Singer, 2017) to investigate potential impact of the increasing complexity of innovation on employee entrepreneurship. Specifically, I link the Census data to USPTO patent data to focus on a near census of highly innovative firms and startups founded by entrepreneurs who had been previously employed at those firms in 29 states of the U.S. between 2000 and 2014. I then construct a firm-level measure of the complexity of innovation, which I call "knowledge interdependence", that captures how intertwined various different technologies are in producing innovations.[11] I then address the following research questions. How does knowledge interdependence affect employee entrepreneurship? What type of individuals (employees of potential parent firms) will be affected more than others? And importantly, how does this affect the performance of startups?

In the existing literature, employee entrepreneurship is explained as driven by profitable opportunities that are unexploited by the focal firm (Franco and Filson, 2006; Freeman, 1986; Hellmann, 2007; Klepper and Thompson, 2010; Gambardella et al., 2015). While various labor market frictions in this regard have been examined in the literature (Campbell et al., 2017), the one of particular interest here is the technological knowledge interdependence in highly innovative firms. It has been shown, in the context of some particular industries, that innovative employees possessing complex knowledge are less likely to move and become employee entrepreneurs (Ganco, 2013) as complex knowledge is difficult and costly to transfer (Rivkin, 2000; Rivkin and Siggelkow, 2002; Grant, 1996; Nickerson and Zenger, 2004). However, to the best of my knowledge, no extant study has looked at the same for the whole U.S. high-tech sector. More importantly, while each knowledge component could be very complex or simple in itself, the extent to which those knowledge components are interconnected to each other is likely to be a strategic choice of firms but not individual employees. As a

[11] See below for the precise definition of "highly innovative firms" and the construction of the measure of complexity of innovation.

result, the knowledge components involved in firm innovation activities could be either integrated and interdependent with each other or separated into loosely coupled modules or clusters.

The empirical results suggest that knowledge interdependence in the incumbent firm is in fact negatively associated with the likelihood of entrepreneurial entry. Moreover, startups founded by previous employees of parents with high knowledge interdependence are more likely to survive. What is the most plausible mechanism that is driving these relationships? One may conjecture that increasing knowledge interdependence may lead to a reduction in employee entrepreneurship simply by raising the bar for entry, which lowers the likelihood of transition into entrepreneurship but has a positive selection on startup quality. That is, given that spinouts tend to explore ideas closely related to those pursued by their parent firms (Golman and Klepper, 2016), potential employee entrepreneurs might face higher uncertainty and incur higher costs to mobilize or/and re-create various knowledge components that are closely integrated and highly complementary to each other (Nickerson and Zenger, 2004; Campbell et al., 2012). If knowledge interdependence simply raises the bar for entry, however, we would expect that individuals (potential entrepreneurs) having superior human capital and capabilities would be more likely to clear this bar so that the negative effect on the selection into entrepreneurship would be less pronounced in the right tail of human capital distribution. Indeed, existing literature finds that individuals having higher human capital are less likely to leave but more likely to form a startup if they do leave their current employer (Campbell et al., 2012). As shown below, however, empirical evidence is not in accordance with this line of reasoning.

Instead, I find the opposite, namely, that the negative impact on entry and the positive selection effect on startup quality is significantly more, not less, pronounced among high-performing and high-ranking individuals.

This puzzle suggests that an alternative mechanism via the internal labor market could be at play inside incumbent firms. Specifically, since high-performing individuals at the right tail of their firm's earnings distribution are more likely to have access to and control of important complementary assets, firms are willing to provide them with ample incentive not to leave especially when losing such human capital has a detrimental effect on firm performance (Carnahan et al., 2012; Campbell et al., 2012). In particular, as knowledge interdependence creates stronger complementarity and needs for collaborative problem-solving (Nickerson and Zenger, 2004), parent firms relying on more interconnected knowledge networks would be even more willing to offer better pay-for-performance incentives to attract and retain valuable human assets. Thus, knowledge interdependence does not only affect employee entrepreneurship in a simple way by causing labor market frictions in preventing mobility (Campbell et al., 2017), but it also demands proactive changes in managerial practices.

Indeed, while startups founded by former employees of firms relying on higher degree of knowledge interdependence are more likely to survive, the selection effect is even stronger among startups founded by individuals ranked higher in the human capital distribution. This fact further falsifies the alternative explanation that knowledge interdependence in parent firms merely reduces the likelihood of employee entrepreneurship by imposing higher costs of entry and mobility. Since if it were the mechanism, individuals who are more capable of mobilizing and/or creating complementary capabilities to start their own businesses would be less affected than less capable ones.

Together, the findings in this chapter provide important implications for big economy-wide issues. First, this study contributes to the existing literature on declining business dynamics by showing that growing knowledge interdependence inside highly innovative firms could be a potential reason why firm entry rate in the U.S. high-tech sectors has been declining in recent decades (Haltiwanger et al., 2016). It also raises an important question to be fully examined. That is, as innovation becomes more and more complex, do large established firms start to substitute for the role of start-ups in pushing forward the technological frontier and driving economic growth? As organizational economists have long maintained that efficient knowledge hierarchy could better deal with complex knowledge needed in the production process (Garicano, 2000; Garicano and Rossi-Hansberg, 2004), allocating complex innovation activities to large established firms might thus be efficient, especially when most start-ups are inevitably limited in size to achieve the efficient knowledge hierarchy. If this is the case, declining business dynamism, at least in the high-tech sector, may just be a reflection of technological change and efficient (re)allocation of resources but not necessarily detrimental to technological advancement and economic growth. That said, it has also been found that larger firms may be simply better at capturing value from innovation, not producing innovations of higher quality (Arora et al., 2022). The importance of a further, in-depth examination of these issues thus cannot be overestimated.

Also, this study contributes to the strategic human capital and employee entrepreneurship literature by bridging existing studies examining the role of knowledge complexity, complementary assets, and compensation dispersion on employee entrepreneurship (Ganco, 2013; Campbell et al., 2012; Carnahan et al., 2012). In particular, my findings here indicate that knowledge interdependence does not simply raise the entry costs and uncertainty of entrepreneurship, but incumbent firms are proactively managing their valuable human assets to cope with the increasing knowledge interdependence required for innovations. Indeed, it is possible that incumbent firms with higher knowledge interdependence uses better pay-for-performance compensation schemes to retain valuable human assets to sustain their technological advantages (Carnahan et al., 2012). In its turn, this may suggest that the growing knowledge interdependence as an important aspect of technological change could potentially be an important source of rising income inequality in recent decades (Acemoglu and Autor, 2011; Song et al., 2019), at least in the high-tech sector, which I formally examine in Chapter 3.

This chapter is organized in a way to reflect the process of abductive reasoning.[12] In the next section, I describe data sources and methodologies to construct the main sample and key variables of interest. Then, I present some facts on the time trend of the growing firm knowledge interdependence and declining employee entrepreneurship. Next, to probe the relationship between knowledge interdependence and employee entrepreneurship and the most plausible explanation of this

[12] While it has long been recognized as one of the three basic forms of human reasoning (Peirce, 1997), abductive reasoning has recently been rediscovered in the strategy, organization, and economic literature (Pillai et al., 2021; Shani et al., 2020; Heckman and Singer, 2017; Tohmé and Crespo, 2013). In essence, in search of the most plausible explanation, abductive reasoning "produces exploratory hypotheses" and "yields plausible explanations about puzzling phenomena . . . following a question, such as 'What is going on?' "(Shani et al., 2020).

phenomenon, I investigate the relationship between firm knowledge interdependence and the likelihood of employee entrepreneurship, including heterogeneity with respect to individual human capital. Finally, I examine the survival of startups formed by employee entrepreneurs.

Data

The primary data sources for this study are the USPTO patent data, the Longitudinal Employer-Household Dynamics (LEHD), and the Longitudinal Business Database (LBD) from the U.S. Census Bureau. The LEHD is a matched employer-employee dataset that includes employment history and quarterly wage information collected from the state un- employment insurance (UI) system and the associated ES-202 program. Under most state laws or regulations, wages reported to the state UI system include bonuses, stock options, severance pay, the cash value of meals and lodging, tips, and other gratuities. In some states, wages also include employer contributions to certain deferred compensation plans, such as 401(k) plans.[13] The LEHD also contains individual identifiers and employer identifiers so that job mobility could be tracked over time. In this paper, I use the LEHD employment and wages data from 29 states provided by the U.S. Census Bureau through its Research Data Center (RDC).[14]

The LBD includes all U.S. non-farm establishments with paid employees and contains information on employment, year of entry and exit, payroll, industry, and

[13] See https://www.bls.gov/cew/overview.htm for more details.
[14] The 29 LEHD states are AL, AZ, CA, CO, CT, DE, ID, IN, KS, MD, ME, ND, NE, NJ, NM, NV, NY, OH, OK, OR, PA, SD, TN, TX, UT, VA, WA, WI, and WY.

geographic location. I use the LBD to construct firm-level variables including firm size, age, growth, the degree of industry diversification, and so on. While the Census has provided time-consistent establishment identifiers, the firm identifiers in the current version of the LBD are not perfectly longitudinally consistent (Chow et al., 2021). Since I need to track innovation activities in the same firm over time, I adopt a methodology similar to (Dent et al., 2018) to create a longitudinally consistent identifier for firms in the LBD. Then, I use the internal bridge file provided by the Census at the Employer Identification Number (EIN) and state level to link the LBD to the LEHD.

It is noteworthy that neither the State Employer Identification Number (SEIN) nor the federal Employer Identification Number (EIN) in the LEHD can be used as the appropriate firm identifier since those are based on tax files or the state UI system. More importantly, they do not capture firm organization and governance structure. In other words, a multi-unit enterprise might have multiple SEIN and EIN associated with it even in the same year. The LBD, however, provides a more appropriate firm identifier "lbdfid" that better identifies business establishments over which a firm has operational control. Thus, instead of using SEIN or EIN as the firm identifier, I use a self-constructed longitudinally consistent firm identifier in the LBD and the EIN-SEIN bridge between the LBD and LEHD to group all associated SEIN in the same state to be the same firm.

There are three interconnected samples I used for the analysis in this paper. First, by using the patent assignee-firm crosswalk created by the Census (Dreisigmeyer et al., 2018), I link the LEHD and LBD to the USPTO patent data and track innovation activities inside a near census of patenting firms in the U.S. between 2000 and 2014. Since the focus of this paper is on employee entrepreneurship in the private sector, I further exclude firms and organizations in educational services, finance and insurance, and public administration from the current analysis. Second, I identify all startups founded by previous employees of these highly innovative firms to investigate their performance by using information from the LBD. A detailed description of identifying startups and linking them to parent firms is provided in the following subsection. Third, I collect all the individual-level data including both startup founders and non-founder employees at those highly innovative firms to conduct individual-level analysis in search of the best alternative explanations.[15] Similar to previous studies (Sorkin, 2018; Tate and Yang, 2015; Card et al., 2013; Carnahan et al., 2012), I restrict the sample to individuals with age between 18 and 65 who earn more than $3,000 in 2000's dollar value and use the annual dominant job that has the highest annualized pay for each individual to track job mobility and transition from waged worker to startup founders.

Identifying Employee Entrepreneurs and Parent Firms

To identify startup-parent linkages and startup founders, I begin by constructing a sample of all startups in the U.S. economy between 2000 and 2014 in the LBD. Importantly, I combine the age information of new establishments in the LBD with employment and wage information in the LEHD to enhance the time accuracy of firm entry and eliminate newly established businesses that belong to existing firms

[15]Note that if a patenting firm does not spawn any startups and founders in a given year, it will still be included in all three samples.

(Haltiwanger et al., 2016; Choi et al., 2021). Then, following (Choi et al., 2021), I identify startup founders as individuals who join the startup in the same quarter as the firm entry and rank among top three earners during the first year of startup formation. While this classification may not be able to identify all startup owners perfectly, it is very likely to incorporate the vast majority of startup owners in the data. Indeed, Azoulay et al. (2020) find that around 90% of S corporation owners identified by K-1 filings data also appear as the top three earners during the firms' first year in the W-2 and LEHD data.[16] Since I restrict the sample to annual dominant jobs, this also means that the job at the startup is the dominant job of identified founders.

Next, I use the LEHD to trace the employment history of all startup founders and identify parent firms as their most recent annual dominant employer before those founders joined the startup. In particular, I allow gap years during the transition from waged employees to startup founders in consideration of the possibility of an incubation period of starting a new firm without setting up a payroll. I also use an alternative way of identifying "spinout" and their linkages with parent firms by looking at moving clusters of the entire founding team from incumbent firms to startup (Starr et al., 2018; Sakakibara and Balasubramanian, 2020), and every member of the founding team is identified as an employee entrepreneur. The main results are robust and qualitatively similar across these two very different identification methodologies. In the following analysis, I use the former one as the main methodology to construct the startup-parent linkage and identify startups and their founders.

[16] Since owners of sole proprietors and partnerships are prohibited from paying themselves wages, they do not appear in the LEHD. As a result, they will not be captured by this definition of founders.

Highly Innovative Firms and Knowledge Interdependence

As mentioned, the main focus of this chapter is to examine how knowledge complexity, measured by how distinct knowledge components are interconnected with each other, affects potential entrepreneurial activities originating from highly innovative parent firms. The sample construction starts with the USPTO PatentsView data and matches it to LBD firms by using the Business Dynamics Statistics of Patenting Firms (BDS-PF) crosswalk provided by the Census (Dreisigmeyer et al., 2018). The BDS-PF crosswalk matches more than 90% of all USPTO patents that are granted between 2000 and 2015. This means that the BDS-PF crosswalk also includes patents that were applied for in earlier years. Including those patents could potentially lead to selection bias if innovation activities are measured based on patent application years. For example, a patent application filed in the late 1990s will be omitted from the sample if it was granted before 2000. Thus, I exclude all those patents in the BDS-PF crosswalk with application years earlier than 2000 and use the patent application year as the time of firm innovation activities. Since the last available year of the LEHD is 2014 and job mobility is measured on an annual basis, the sample period is shortened to 2000-2013 after I merge the USPTO patent data to the LEHD and LBD. I also utilize the inventor location information and require that a patent is matched to a firm if at least one of its patent inventors is located in the same state as the firm.

Furthermore, I define "highly innovative firms" as those that had at least two patents applied for and granted by USPTO in two different years during the sample period. Having at least two patents in two different years is necessary to be able to employ firm fixed effects in estimations below but, more importantly, it excludes firms whose innovative (patenting) activity was of a one-off nature so that they can hardly be deemed "highly innovative."

To measure knowledge complexity through technological knowledge interdependence of incumbent firms, I follow the prior works (Fleming and Sorenson, 2001, 2004; Sorenson et al., 2010; Ganco, 2013; Yayavaram and Ahuja, 2008), and identify knowledge interdependence by relying on the co-occurrence of patent technology groups. Specifically, I collect all patents that were applied for by a firm in a given year to construct a technological knowledge network. Each Cooperative Patent Classification (CPC) technology group is a "node" in such a network[17]. And an "edge" connecting a pair of technology components is identified when those two CPC technology groups are co-assigned to the same patent.[18]

The computation of knowledge interdependence involves two steps and is analogous to previous research that focuses on inventor knowledge complexity (Ganco, 2013) and on the complexity of patents (Fleming and Sorenson, 2001, 2004; Sorenson et al., 2010). First, I calculate the frequency of co-occurrence for each CPC technology group in the firm's technology network and compute the interdependence for each knowledge component (technology group) c of firm j in year j as

$$K_{c,j,t} = \sum_{x \in \phi_{j,t}, x \neq c} N^{j}_{c,x,t} / N^{j}_{c,t} \tag{1}$$

[17] Cooperative Patent Classification (CPC) is a bilateral classification system jointly developed by the United States Patent and Trademark Office (USPTO) and the European Patent Office (EPO). CPC is jointly managed and maintained by both offices and is available for public search for classification.
[18] Alternatively, I use a 3-year time window including all patents that are applied by the firm in year t, t 1, and t 2 as an alternative way to construct the firm's technology network. The main results remain qualitatively similar.

where $\phi_{j,t}$ is the set of all technology groups in firm j, N_j is the number of patents co-assigned to CPC group c and x, and N_j is the number of patents assigned to CPC group c. Then, the knowledge interdependence of firm j in year t is measured as the average of $K_{c,j,t}$ over the number of technology groups $N_{j,t}$ in $\phi_{j,t}$.

$$K_{j,t} = \sum_{c \in \phi_{j,t}} K_{c,j,t} / N_{j,t} \tag{2}$$

In essence, $K_{j,t}$ measures the expected number of technology groups combined with a randomly selected technology group in the firm j's patent portfolio in year t. Thus, it measures the extent to which a given knowledge component is "interdependent" with other knowledge components in the firm's technology network. Note that since it is impossible to measure interdependence in patent portfolios comprised of only one CPC technology group, the above procedure restricts the sample of firm-year observations to those having more than one CPC technology group. Thus, the final sample of "highly innovative firms" used for the empirical analysis in this paper includes firms whose knowledge base also meets the minimum complexity criterion, namely, those that have patent portfolios comprised of at least two distinct technological groups.

Firm and Individual Control Variables

As knowledge interdependence is not randomly endowed to incumbent firms, I include a list of variables controlling for characteristics of the parent firms, startup founders, and their coworkers to minimize the risk of potential confounding effects.

Specifically, I include the log number of employees to control for firm size[19] since large firms are found to be less likely to trade their patented knowledge than small firms (Figueroa and Serrano 2013); hence, such firms may have more idle knowledge that could be exploited by potential employee entrepreneurs. Firm age is included since potential disagreements could build up and uncertainty of the quality of innovative ideas might be resolved as the firm ages (Klepper and Thompson, 2010). Industry diversification is also found to be associated with fewer startups (Gompers et al., 2005), and firms with a higher degree of diversification have been found to be better at redeploying human capital in their internal labor markets (Tate and Yang, 2015). Finally, employment growth is included to control for the opportunity costs faced by potential employee entrepreneurs.[20] Following Davis et al. (1998) (DHS thereafter), the employment growth rate $g_{i,t}$ is calculated as:[21]

$$g_{i,t} = \frac{E_{i,t}-E_{i,t-1}}{X_{i,t}} \tag{3}$$

where $X_{i,t} = \frac{1}{2}\left(E_{i,t} + E_{i,t-1}\right)$, and $E_{i,t}$ is the total employment at firm i and year t.

[Insert Table 7 around here.]

Table 7 provides summary statistics of all variables used at the firm level. It is noteworthy that on average, firms have 6.77 distinct CPC technology groups in their patent portfolio in a certain year, each integrated with 1.73 other CPC technology

---

[19] Alternatively, the LBD also provides information on the total payroll of each firm-year. Using total payroll to measure firm size leads to qualitatively similar results.

[20] The assumption here is that potential startup founders can anticipate the growth prospects of their current employers and will be less likely to leave if they perceive that their current employer is on a high-growth trajectory.

[21] This growth rate measure has become standard in the analysis of the establishment and firm dynamics, because it shares some useful properties of log differences but also accommodates entry and exit (Davis et al., 1998).

groups. The average firm has 158.53 employees with age 20.42 years and DHS employment growth rate of 12.13 percent. Moreover, 15.14 percent of all firm-year observations has been classified as "Young Firm" with age less than 5 years old. In this sample of highly innovative firms with at least two distinct CPC technology groups in their knowledge base, the unconditional probability of spawning at least one startup in the next year is 31 percent, with the average number of startups spawned being 1.03.

[Insert Table 8 around here.]

In the following analysis, I also utilize various individual characteristics of founders and their coworkers at the parent firm including age, job tenure, within-firm earning rankings (the percentile of the individual in their firm's earning distribution), earning residuals (from a Mincerian earnings regression), and (for startup founders) the number of different industries the founder previously worked at as a proxy for the diversity of his/her own previous job experience. Table 8 provides the summary statistics of these individual characteristics for all employees at parent firms and for primary startup founders among them, defined as the top earners during the first year of startup formation. It is noteworthy that the unconditional probability of a current employee becoming a startup founder in a given year is 0.1 percent, which means one out of one thousand employees at the incumbent firm is identified as a startup founder in the next year. Overall, the average individual at parent firms earns $51,021 a year in 2000's dollar value. Also, primary founders on average are younger with shorter job tenure at parent firms but earn more and are ranked higher than all employees in their parent firms, both in the earning distribution hierarchy and in terms of earnings residuals.

Facts about Firm Knowledge Interdependence and Employee Entrepreneurship

Before exploring the relationship between firm knowledge interdependence and employee entrepreneurship, I first investigate if and how knowledge interdependence within the sample of highly innovative firms has been changing over time, and whether the time trend is driven by a specific group of firms. Moreover, I show that the likelihood of transitioning into employee entrepreneurship has been declining over the sample period 2000-2014. Importantly, the decline is significantly more pronounced among individuals at the right tail of human capital distribution.

Growing Knowledge Interdependence Inside Highly Innovative Firms

Existing literature finds that the complexity of innovation has been rising as measured by such proxies as larger inventor team size, greater specialization, and more R&D input per unit of innovative outcome (Bloom et al., 2020; Luo and Wood, 2017; Jones, 2009; Wuchty et al., 2007). While these findings suggest that the complexity of innovation itself has been increasing over time, how the organizational structure of technological knowledge within firms has been evolving is relatively understudied in the literature. To verify if knowledge interdependence in firms is indeed rising over time, I estimate a simple OLS regression model:

$$K_{j,t} = \alpha + \beta T + \lambda X_{j,t} + \delta X_{j,t} \times T + \epsilon_{j,t} \quad (4)$$

where $K_{j,t}$ is the measure of knowledge interdependence in firm j, year t constructed as above; T is the time trend; and $X_{j,t}$ is the vector of firm characteristics including 1) firm size as measured by the log of total employment, 2) firm age and 3) industry diversification measured by one minus the Herfindahl-Hirschman Index (HHI) of employment concentration in 4-digit NAICS industries. Industry and state fixed effects are also included to control the difference in levels of knowledge interdependence across industries and states.

[Insert Table 9 around here.]

Table 9 shows the results. Column (1) shows that the average knowledge interdependence among patenting firms has been rising over time. The coefficient estimate on time trend implies that by the end of the sample period in 2013, the average firm knowledge interdependence was approximately 23% higher than it was in 2000. More interestingly, as indicated by Columns (2) and (4), while the levels of knowledge interdependence in large and diversified firms are generally lower than in other firms, the time trends are indistinguishable and ubiquitous as suggested by the insignificant coefficients of the interaction terms. The time trend, however, is distinct across firms of different ages.[22] In particular, Column (3) shows that the increase in knowledge interdependence has been even more pronounced in startups – firms less than five years old. Column (5) includes all the variables of interest at the same time, and similar qualitative results hold. In particular, after controlling firm size and industry

[22] While no figures can be shown due to disclosure requirements at this stage, the time trend of increasing knowledge interdependence and the narrowing trend between young and old firms are indeed visibly prominent.

diversification, knowledge interdependence in startups is indistinguishable from established firms in 2000. In 2013, however, knowledge interdependence in startups is on average 10% higher than in established firms at the mean.

These results suggest that there is a ubiquitous increase in knowledge interdependence in all firms in the sample between 2000-2013, while it is even more pronounced both statistically and economically among startups. The latter implies a possible selection process on new firm formation as the overall knowledge interdependence has been rising in the most innovative sectors of the private economy in the U.S. I will come back to this selection issue in subsequent sections.

Declining Employee Entrepreneurship

Are employees of highly innovative firms becoming less and less likely to leave and found their own business ventures? As the U.S. has experienced a generally declining trend in firm entry rates, especially in high-tech sectors (Decker et al., 2014; Akcigit and Ates, 2019; Haltiwanger et al., 2016), it is important to know if and which type of potential entrepreneurs employed in highly innovative firms are becoming less likely to choose entrepreneurship over time.

For each firm in year t, I construct a dummy variable indicating whether any employees left to start a new firm as the main outcome variable of interest at the firm level. Then, I estimate a simple univariate linear probability model by regressing this dummy indicator for employee entrepreneurship on a time trend variable T:[23]

[23] The regression approach is used here instead of figures showing the time trend because I cannot disclose any figures using the Census data at this stage.

$$Y_{j,t+1} = \alpha + \beta T + d_i + \epsilon_{j,t} \tag{5}$$

where d_i is either the industry-state fixed effect or the firm fixed effect. To further investigate which type of potential entrepreneurs are more or less likely to start their own ventures over time, I then estimate the following regression:

$$Y_{i,j,t+1} = \alpha + \beta_1 T + \beta_2 T \times K_{j,t} + \lambda Z_{i,t} + D_{j\times t} + \epsilon_{i,j,t} \tag{6}$$

where $Y_{i,j,t+1}$ equals one if individual i employed in firm j leaves and is observed as a founder of a new firm in the next year; $K_{j,t}$ is the knowledge interdependence in firm j in year t; $Z_{i,t}$ are individual controls including age, job tenure, proxies for human capital[24], earnings rankings, gender, and race dummies; $D_{j\times t}$ is the firm-year fixed effects that absorb all time-variant and time-invariant firm-level factors, e.g., firm size, age, knowledge interdependence, and so on.

[Insert Table 10 around here.]

As shown in Column 1, Table 10, the likelihood of employee entrepreneurship at the firm level has been declining over time on average by 1.2% each year when the state-industry fixed effect is included. This effect is both statistically and economically significant given the overall average probability of employee entrepreneurship at the firm level in this sample is around 31%. A similar result holds if the firm fixed effect is included to isolate between-firm variations in the sample (Column 2). More importantly, as shown in Columns 3-5, the decline in the likelihood of transition into entrepreneurship is much more pronounced among individuals with high human capital regardless of which proxy for human capital is used. Notably, as shown in Column 4,

[24] See details on the construction of these proxies in the next section.

an individual ranked among the top 5% in the within-firm income distribution is more likely to become a founder than others by 0.28% in 2000.[25] By the end of the sample period in 2013, however, the difference between top 5% earners and other employees in their likelihood of becoming a founder is narrowed to 0.13%. As founders with higher human capital are more likely to found high-growth high-opportunity startups, the declining entrepreneurial activities particularly among high-performing individuals could have a huge impact on economic growth even though it might benefit incumbent firms to sustain competitive advantage.

In sum, this empirical evidence suggests two important and prominent trends inside the highest innovative firms in the U.S. economy – the growing knowledge interdependence and the declining rate of employee entrepreneurship. Are these two trends merely coincidence? If not, how does knowledge interdependence affect individual decisions to become entrepreneurs? I now turn to the multi-level analysis to answer this question in the following sections.

The Likelihood of Employee Entrepreneurship

The main interest of this study is how knowledge interdependence affects employee entrepreneurship and startups founded by previous employees of highly innovative firms. Moreover, the LEHD provides a unique opportunity to investigate potential heterogeneous effects of knowledge interdependence on individual

[25] Note that on average the likelihood for an individual to become a startup founder is 0.1% in the sample.

employees with different human capital endowments. In this section, I describe the empirical design and present the main results at the firm level that firm knowledge interdependence is negatively associated with employee entrepreneurship. Then, I utilize individual-level data from the LEHD and show who in the incumbent firms is more likely to be affected by higher knowledge interdependence.

Empirical Approach

As a first step, I investigate the relationship between knowledge interdependence and the likelihood of startup formation in highly innovative firms. For each firm in year *t*, I construct a dummy variable indicating whether the firm spawns any startups in year *t+1* as the main outcome variable of interest at the firm level. Then, I estimate a linear probability (OLS) model by regressing this dummy indicator for startup formation in year *t+1* on the knowledge interdependence $K_{j,t}$ of firm *j* in year *t*:

$$D_{j,t+1} = \alpha + \beta K_{j,t} + \lambda Z_{j,t} + \delta_t \times \gamma_{state} \times \nu_{industry} + d_j + \epsilon_{j,t} \qquad (7)$$

where $Z_{j,t}$ is a set of firm control variables including the number of technology groups in the firm's patent portfolio, firm size, industry diversification, employment growth, and firm age; $\delta_t \times \gamma_{state} \times \nu_{industry}$ is the year-industry-state fixed effect and d_j is the firm fixed effect (included in some specifications below). These firm control variables are included to isolate potential endogenous effects, and standard errors are clustered at the firm level to allow inter-temporal correlation within the same firm due to unobserved variables.

I also estimate several alternative model specifications for robustness checks. First, since knowledge interdependence is naturally correlated with the scope of the firm knowledge base, i.e., the number of technology groups, multicollinearity could bias the coefficient estimates if there are correlated omitted variables. To check if there is this issue, I use the number of technology groups as a categorical variable instead of a continuous variable in Equation (7).[26] By doing so, the linear correlation between knowledge interdependence and the number of technology groups is eliminated such that the source of identification is by comparing knowledge networks that are identical in scope and size. Furthermore, I estimate Equation (7) by using logistic regression. Given the large number of firms in the sample, it is not computationally feasible to include firm fixed effects. Nevertheless, I include year, state, and industry fixed effects to control for potential unobservable factors.

Furthermore, I use the IHS transformed number of startups $Y_{j,t+1}$ spawned by firm *j* in year *t+1* as an alternative outcome variable to check the consistency of the main results.[27] That is, if higher knowledge interdependence is associated with lower likelihood of startup formation, we would also expect to see fewer number of startups originating from parent firms with higher knowledge interdependence. Similar as before, I estimate the model by OLS including the year-industry-state fixed effect and firm fixed effect as well as the negative binomial model with year, industry, and state effects. The main results are presented in the next section.

---

[26] Using the number of technology groups as a categorical variable requires that there are more than one firm-year observations in each category. This is rarely the case for firms having more than 95 technology groups in the data. To ensure that the regression analysis is estimated based on the same sample, I group all firm-year observations with more than 95 technology groups into the same category for this specific model specification. Excluding those observations lead to qualitatively similar results.

[27] IHS (inverse hyperbolic sine) transformation is used for the number of startups because of zeroes in observations. It is defined as $y = \ln{(x + \sqrt{x^2 + 1})}$. The IHS transformation approximates the log transformation much better than $y = \ln{(1 + x)}$ for small values of *x*.

Main Results

Table 11 presents results of estimating Equation (7) which examines the relationship between startup formation and knowledge interdependence at the firm level. Column (1) shows that higher knowledge interdependence in a given firm is associated with lower likelihood of generating startups, and the relationship is both statistically and economically significant. More specifically, one standard deviation in knowledge interdependence (which is 0.521 – see Table 7 above) is associated with about 2.2% percentage points decrease in the probability of startup formation; that is, it decreases the probability of spawning startups by about 7% for an average firm whose probability of having at least one startup in year *t+1* is around 31% (Table 7 above). Column (2) addresses the potential multicollinearity issue by using the number of technology groups as a categorical variable. The magnitude of the coefficient on (logged) knowledge interdependence is somewhat lower, but it remains highly significant, both economically and statistically. In Column (3), I include firm fixed effect to control for potential time invariant unobservable at the firm level. While the negative relationship between knowledge interdependence and the propensity of startup formation still holds, both statistical and economic significance levels are much lower. This is not surprising since much of the variation of knowledge interdependence comes from between-firm differences. By construction, a patenting firm needs to have patent applications in multiple years to make it possible to create sufficient within-firm variation in its patent portfolio over time. Columns (4) and (5) present the results from estimating logistic and probit regression models, respectively. The results remain qualitatively similar.

[Insert Table 11 around here.]

The relationship of the control variables to the likelihood of startup formation is all consistent across different model specifications and is mostly consistent with existing literature. The number of technology groups in the firm's patent portfolio is positively associated with the probability of startup formation, which indicates that firms having higher innovative capability are more likely to spawn new ventures. Also consistent with previous findings (Klepper and Thompson, 2010), older firms are less likely to spawn startups. Moreover, as conjectured, higher firm growth is significantly associated with fewer startups as employees at firms with high growth potential face larger opportunity costs than employees in low-growing firms to leave their current jobs to start their own ventures. In contrast to what previous literature has found (Gompers et al., 2005), however, more diversified firms are associated with more startups, not less.

[Insert Table 12 around here.]

Column (1) to (3) in Table 12 shows the results of the same estimations but using the IHS-transformed number of startups as an alternative outcome variable. In Column (4), I estimate a negative binomial model with the outcome variable as the number of startups in year *t+1*. Similar to Table 11, the coefficient on the knowledge interdependence variable remains negative and statistically significant, and the signs of all control variables remain the same as before. It is noteworthy that the significance level with firm fixed effect is improved in this model specification in Column (3). This is expected since there is more variation in the outcome variable when the dummy indicator for startups is replaced by a continuous variable counting the number of startups.

Heterogeneous Effects on High-performing Individuals

Why and how does higher knowledge interdependence reduce the likelihood of transition into entrepreneurship? To answer this central question, it is essential to open the "black box" of the firm. In this section, I utilize detailed data including all employees of the patenting firms in the LEHD to explore plausible explanations. Specifically, I ask the following questions. Who are the founders of startups originating from highly innovative firms? If higher knowledge interdependence reduces the likelihood of startup formation, does it affect a specific group of employees of incumbent firms, e.g., those with higher human capital endowments, more than others?

I utilize two alternative ways of measuring the human capital of employees in incumbent firms – earning residuals and earning rankings. First, I estimate earnings residuals for all individuals employed at incumbent firms by the following (Mincer-type) equation:

$$Log(w_{i,t}) = \beta_0 + \beta_1 X_{i,t} + \delta_t \times d_j + \eta_{i,t} \tag{8}$$

where $X_{i,t}$ is a group of individual characteristics including age, age squared, gender, race, education background, current job tenure, and the number of industries the individual had previously worked in (as a proxy for the variety of past job experience). The specification also includes the firm-year fixed effect $\delta_t \times d_j$, so the residual $\hat{\eta}_{i,t}$ from this regression could serve as a proxy for individual's human capital. In other words, the earning residuals from this model specification can be interpreted as the relative human capital among all individuals employed at the same firm in a given year.

As an alternative (and more straightforward) measure, I use the individual's percentile in his or her employer earnings distribution in a given firm-year (earnings ranking, hereafter) as a proxy for their position in the hierarchy within the firm's job ladder. Since LEHD covers all individuals employed by the firm, their earning rankings could serve as an indicator of job seniority. Indeed, previous studies have used earning rankings to identify top individuals as well as managers at each business establishment (Braguinsky et al., 2014; Tate and Yang, 2015).

To explore potential heterogeneous effects of knowledge interdependence, I estimate the following linear probability model at the individual level:

$$D_{i,j,t+1} = \beta_0 + \beta_1 K_{j,t} \times \hat{\eta}_{i,t} + \beta_2 \hat{\eta}_{i,t} + \gamma X_{i,t} + \delta_t \times d_j + \mu_{j,t} \qquad (9)$$

where $D_{i,j,t+1}$ is a dummy variable indicating whether an individual *i* becomes a startup founder in year *t+1*; $K_{j,t}$ is the knowledge interdependence of firm *j* in year *t*; $\hat{\eta}_{i,t}$ is the earning residuals estimated by Equation (8); $X_{i,t}$ is a set of individual controls including age and its quadratic term, current job tenure and its quadratic term, gender and race; and δ_t and d_j are year and firm fixed effects, respectively. In an alternative specification, I replace $\hat{\eta}_{i,t}$ with the individual's earning rankings, as discussed above. Notice that the firm-year fixed effect is included to control for all time-invariant and time-variant factors at the firm level. In particular, the main effect of knowledge interdependence, along with all other firm-level characteristics in each year is absorbed by the firm-year fixed effect. If firm knowledge interdependence has a heterogeneous effect on individuals having different human capital, we should expect β_1 to be significantly different from zero. The estimation results are shown in Table 13. The key finding is that across all specifications, higher firm knowledge interdependence has significantly stronger negative effects on the likelihood of becoming a founder among high-performing and highly ranked individuals of the incumbent firm.

[Insert Table 13 around here.]

As shown in Column (1), while employees with high earning residuals are on average much more likely to become founders than employees with low earnings residuals, the interaction effect of knowledge interdependence and earning residuals is negative and significant.[28] The marginal effect of earning residuals at firms with mean knowledge interdependence is 0.49×10^{-3}. Thus, one standard deviation increase in earning residuals at firms with mean knowledge interdependence would on average increase the probability of becoming a founder by 0.026 percent, or about 26 percent of the unconditional probability of becoming a founder (0.1 percent as shown in Table 8). Given this, the economic impact of the interaction term is indeed large. Specifically, one standard deviation increase in knowledge interdependence is estimated to reduce the marginal effect of earning residuals by 0.0065 percent, or about 9.7 percent of the main marginal effect of earning residuals on the probability of becoming a founder, which is 0.067 percent.

In Column (2), I use earning rankings as an alternative measure of the relative human capital of individuals. Consistent with the results in Column (1), while higher-

[28] Note, once again, that the effect of knowledge interdependence itself is absorbed by the firm-year dummy.

ranked individuals are more likely to found startups, the negative impact of higher knowledge interdependence is also significantly stronger among those individuals. Indeed, this negative heterogeneous effect is further confirmed when I replace earning rankings with an indicator variable for the top five percent earners who are presumably individuals in the top management layers. As shown in Column (3), while the average probability of becoming a founder among the top 5% earners is 0.2 percent higher than other employees at firms with mean knowledge interdependence, one standard deviation increase in firm knowledge interdependence would reduce the probability by 0.017 percent, which is approximately 8.5 percent lower than the baseline effect. Given the unconditional probability of becoming a founder is 0.1 percent, these effects are indeed economically significant.

Additionally, coefficient estimates on individual control variables are qualitatively similar across all specifications. Higher residual earners, those who are ranked higher in firms, male and white, are more likely to become entrepreneurs. The effect of age has an inverted-U shape in Column (1), which is consistent with existing theories that emphasize the role of learning and the quality of ideas (Klepper and Thompson, 2010; Braguinsky et al., 2012).[29] The empirical results presented above seem to contradict a simple interpretation of a lower probability of transition into entrepreneurship from firms with high knowledge interdependence as just being driven by higher barriers to exploiting ideas learned at the previous employer in one's own venture. If that were indeed the case, wouldn't individuals with relatively high human

[29] In columns (2) and (3) the coefficients on age and its squared term are generally small and insignificant, probably due to the correlation between age and the position in a firm's hierarchy.

capital be more likely to be able to overcome such barriers? That seems to imply that we should have observed a positive impact of the interaction between knowledge interdependence and human capital on the probability of forming a startup in Table 13 but what we observed was directly the opposite.

An alternative story starts from the conjecture that high residual earners are more likely to be those who have developed strong complementary assets in the firm. This is further confirmed by using earning rankings as a proxy for job positions within each firm. Intuitively, higher-ranked employees are more likely to be in key positions with greater control of and better access to complementary assets. The fact that those are also the employees who are less, not more likely than lower-positioned employees to leave firms with highly interdependent knowledge to start startups suggests that firms may be utilizing internal labor markets to keep them from leaving. In other words, higher knowledge interdependence involved in firm innovation activities creates strong complementarities between the firm and especially its high-performing and highly ranked employees. As a result, incumbent firms strive to keep them from leaving as losing such employees could have large negative consequences for firm performance. While there could be various managerial practices used by incumbent firms to keep such valuable human assets from leaving, previous study has shown that higher compensation dispersion within firms could be one effective way to do it (Carnahan et al., 2012). I examine this possibility in Chapter 3 of this dissertation. In the next section, I'll further examine the selection effect of higher knowledge interdependence on the quality of startups being formed to falsify the alternative explanation that knowledge interdependence *merely* rises the costs of entry.

Spinouts Survival

The results in the previous section suggest that higher knowledge interdependence reduces employee entrepreneurship in particular among high-performing and high-ranking individuals. Since new business ventures involve a lot of uncertainty, employed workers will choose to become an entrepreneur only if they receive a strong enough good signal about the quality of their underlying business ideas (Braguinsky et al., 2012; Takii, 2003). If higher knowledge interdependence imposes higher costs of entry or frictions on the labor market, potential founders in firms with higher knowledge interdependence will only leave to become entrepreneurs if they receive a stronger positive signal about the probability of success compared to potential founders in firms with less knowledge interdependence. If the signals are correct on average, this selection effect should lead to startups spawned by firms with higher knowledge interdependence to have better performance and higher chances of survival. Moreover, if disproportionately high incentives are indeed provided to high-performing and highly ranked individuals in firms with great knowledge interdependence, we should expect the positive effect of knowledge interdependence on startup survival to be even stronger if the new firm is founded by high-performing and highly ranked employees of their parent firm.

I use the LBD to collect information on the year of exit for all startups founded by employee entrepreneurs included in the previous sample. The LBD provides a time-consistent identifier for business establishments between 1991 and 2018. If there is a change in ownership of a business establishment, e.g., by mergers and acquisitions, it will be assigned a new firm identifier while the establishment identifier would remain the same in the data. Thus, I rely on establishment-level observations in the LBD and consider startup failure when the new firm's physical establishment becomes inactive, e.g., with zero employment, or exits as indicated in the LBD.

By construction, a startup can have up to three founders coming from different parent firms. I use individual characteristics of the founder with the highest earnings during the first year of the startup (the "primary founder" in Table 8 above) to investigate potential interaction effect between the founder's human capital and parent firm knowledge interdependence on startup survival. This ensures that a startup is linked to only one and most important parent firm and individual founder in the following analysis.

To empirically test the above conjecture, I implement the Cox proportional hazards model:

$$\log\left(\frac{h(t)}{h_0(t)}\right) = \beta_1 K + \beta_2 K \times HC + \gamma X + \lambda Z + \Delta + \mu \qquad (10)$$

where the outcome variable is the log of the hazard ratio between *h(t)* – the hazard rate of startup failure in year *t*, and $h_0(t)$ – the baseline hazard; *K* is the degree of knowledge interdependence of the parent firm at the time the founder left to found his/her startup; *HC* is the founder's human capital as measured by earning residuals, earning rankings or a dummy indicator for the top five percent earners; *X* is a group of founder characteristics including age and its squared term, job tenure and its squared term, number of industries previously worked in, gender and race; *Z* is a group of parent firm characteristics including firm size, age, degree of diversification and the number of technology groups; Δ represents a group of dummy variables of year, state, parent industry, and startup industry fixed effects; and μ is the error term. Table 14 presents the results.

[Insert Table 14 around here.]

Column (1) shows the baseline effect of higher knowledge interdependence in parent firms on startup survival. Consistent with the argument above, startups originating from parent firms with higher knowledge interdependence and founded by higher-ranked individuals are more likely to survive. Specifically, one unit increase in the log of knowledge interdependence of a parent firm would lower the hazard of exit by 6.95 percent for startups founded by its previous employees.[30] The magnitude of this main effect is indeed very similar regardless of the measure of human capital used in Columns (3) and (5). Columns (2), (4) and (6) add the interaction term of knowledge interdependence of parent firms and human capital measures of the dominate startup founder. Results consistently show that the negative effect of higher knowledge interdependence in parent firms on the logged hazard ratio is significantly stronger among startups founded by individuals with higher human capital. This suggests that the selection effect of higher knowledge interdependence on startup quality is stronger among those founded by higher ranked individuals. It also implies that knowledge interdependence does not merely impose a higher barrier uniformly to all potential

[30] The estimated ratio of hazards is simply the exponentiated coefficient: $\exp(-0.072) \approx 0.9305$.

entrepreneurs, but it also raises the opportunity cost for employee entrepreneurship in particular among highly ranked individuals.

Some of the control variables are also noteworthy. First, startups founded by previous employees of firms with a broader scope of knowledge, e.g., more technology groups, are more likely to exit. It is possible that while being embedded in an environment with various knowledge pieces facilitates idea generation and thus startup formation, it lowers the bar of employee entrepreneurship so that startups from those firms are on average of relatively lower quality and hence less likely to survive. Second, while the parent firm size and degree of industry diversification do not have a significant impact on startup survival, startups spawned by older firms are more likely to survive. Regarding founder characteristics, there is an inverted U-shaped relationship between startup survival and founder's age and job tenure, which indicates highest chance of survival when the startup founder is of middle age and job experience. Moreover, higher chance of startup survival is associated with a broader industry experience as measured by the number of industries the founder previously worked in.

Overall, these findings provide supporting evidence that there is indeed a selection on the quality of startups by higher knowledge interdependence within parent firms. And this selection effect is especially prominent among potential entrepreneurs who are high performers or highly ranked in the incumbent firm. The latter further confirms that high performers and highly ranked individuals in firms with higher knowledge interdependence are indeed facing higher opportunity costs to quit their paid jobs.

Conclusion

This chapter investigates an important aspect of technological change and the potential impact on employee entrepreneurship. As there has been a decline in U.S. business dynamism, especially in the high-tech sector, it is a cause for concern if there is a fundamental change in the American innovation and entrepreneurship paradigm. My findings suggest that the increasing interdependence of knowledge involved in firm innovation activities is associated with a lower likelihood of employee spinouts. And the relationship is significantly stronger among high-performing and highly ranked employees who are very likely to possess important complementary human assets for the firm. Moreover, startups founded by former employees of firms with higher knowledge interdependence are more likely to survive, and this relationship is significantly more pronounced if the founder ranks at the high-end on the human capital distribution.

Importantly, these empirical findings rule out an alternative explanation that higher knowledge interdependence merely raises the entry barrier into entrepreneurship. If that was so, potential founders with higher human capital should be less, not more, affected by the additionally imposed costs of higher knowledge interdependence to mobilize and/or re-create complementary capabilities. The heterogeneity effects uncovered in the above analysis suggests that higher knowledge interdependence not merely increases labor market frictions. Instead, the most innovative firms may be proactively adjusting human capital management practices, e.g., better pay-for-performance, to retain the most valuable human assets within the firm. I further investigate the plausibility of this alternative explanation in the next Chapter.

Chapter 3: Knowledge Interdependence, Pay-for-Performance, and Income Inequality[31]

Introduction

What is the mechanism through which knowledge interdependence reduces employee entrepreneurship in the most innovative firms? Why is the selection effect stronger among individuals with higher human capital and ranked at the top of the human capital distribution? Empirical evidence in the previous chapter suggests that higher knowledge interdependence does *not* merely impose higher barriers to entry, e.g., by raising the costs of knowledge transfer and mobilization of complementary assets. Instead, firms seem to be proactively managing the most valuable human assets to prevent them from leaving. In this chapter, I seek to further investigate the mechanism underlying the relationship between knowledge interdependence and employee entrepreneurship.

As compensation dispersion is found to be effective tools attract and retain valuable human assets (Carnahan et al., 2012), it is possible that innovative firms relying on interdependent knowledge inputs are utilizing better pay-for-performance compensation scheme to prevent talented works from leaving. However, why do firms pay more if higher knowledge interdependence imposes extra costs of mobility and

[31] This research uses data from the Census Bureau's Longitudinal Employer-Household Dynamics Program, which was partially supported by the following National Science Foundation Grants SES-9978093, SES-0339191 and ITR-0427889; National Institute on Aging Grant AG018854; and grants from the Alfred P. Sloan Foundation. Any views expressed are those of the author and not those of the U.S. Census Bureau. The Census Bureau's Disclosure Review Board and Disclosure Avoidance Officers have reviewed this information product for unauthorized disclosure of confidential information and have approved the disclosure avoidance practices applied to this release. This research was performed at a Federal Statistical Research Data Center under FSRDC Project Number 2296. (CBDRB-FY22-P2296-R9700)

creates more labor market frictions? Indeed, it is plausible that firms with higher knowledge interdependence may appropriate rents from their employees as the employees' value of outside option is now lower. Note, however, that such a conjecture would be inconsistent with the findings in Chapter 2 to the effect that a significantly stronger association between higher knowledge interdependence and less entrepreneurial activity is found among those with higher human capital and/or ranked higher in the job ladder within parent firms. That is, if capable employees ranked higher, e.g., managers, are better at mobilizing and/or re-creating complementary capabilities, the value of their outside option would remain less affected relative to those of less capable employees ranked lower in the focal firm (Campbell et al., 2012; Agarwal, et al., 2016). As a result, their likelihood of transition into entrepreneurship should be less affected by higher knowledge interdependence. Thus, there should be alternative and better explanation(s) to reconcile with these empirical findings.

The value of outside option affects the bargaining power of employees who might be considering leaving the firm to start their own business. This may indeed be lower as knowledge interdependence grows inside the parent firm. At the same time, higher knowledge interdependence can be increasing the value of especially high-level employees to the parent firm, increasing their bargaining power and offsetting the impact of higher costs of transition to entrepreneurship. If the importance of employees with high-level human capital to the parent firm is increasing by more than the costs of transition to entrepreneurship for such employees, firms with higher knowledge interdependence would in equilibrium pay more, *not less*, to their best talents. Below, I use this intuition to develop a formal theoretical model. One implication of the model is that top-level talent will actually be paid more, both in absolute and relative terms, to induce it to stay with the parent firm. In its turn, this drives up both within- and across-firm income inequality over time as innovation becomes more and more complex requiring interdependent knowledge inputs. This alternative explanation is consistent with the hypothesis advanced in the literature, that the persistence in firm pay premiums reflects rent-sharing (Dickens & Katz, 1986; Katz & Summers, 1989). In particular, if complex innovation relying on more interdependent knowledge creates large enough "entrepreneurial profits" (Schumpeter, 1911), firms might be incentivized to share rents with their employees especially with those who are in top positions with control of and access to crucial complementary assets.

In the following sections, I first establish a formal economic model of employee entrepreneurship to explicitly show that this alternative explanation could better explain and reconcile all empirical findings. In particular, I use the LEHD data with individual earnings information to show firms with higher knowledge interdependence are indeed implementing better pay-for-performance, which in its turn creates greater within-firm compensation dispersion as compared to those with lower knowledge interdependence. Importantly, the formal economic model generates hypothesis that firms with higher knowledge interdependence should also pay higher earning premium, which can be tested by future research.

My findings in this chapter speak to the increasing debates on rising wage inequality in the United States since the 1970s (Acemoglu & Autor, 2011). Existing studies have shown that a large share of this rise in wage inequality can be attributed to the rise in the dispersion of average earnings between firms, across industries (Barth et al., 2016; Bernard & Jensen, 1995) and/or occupations (Autor et al., 2008; Autor et al., 2003). Nevertheless, while the differences in firm pay premiums[32] contribute substantially to the distribution of earnings in the U.S (Abowd et al., 2002; Abowd et al., 1999; Goux & Maurin, 1999), the *variance* of firm pay premiums has been approximately stable in the U.S. (Song et al., 2019). Importantly, even though the rise in "between-firm" inequality – the increasing differences in firm average earnings – could be completely explained by changes in the composition of workers between firms (Song et al., 2019), those only account for about two thirds of the rise of earnings inequality in the U.S. economy overall. The remaining one third of such rise in wage inequality, coming from within-firm increase in earnings dispersion remains understudied.

Existing studies attribute the rise of within-firm inequality to the rise in domestic outsourcing, temporary work (Abraham & Taylor, 1996; Goldschmidt & Schmieder, 2017; Katz & Krueger, 2019; Segal & Sullivan, 1997; Weil, 2014), and technological changes that shift the demand for different skill groups (Acemoglu & Autor, 2011; Song et al., 2019). I contribute to this conversation by proposing an alternative mechanism that could potentially drive up income inequality in particular *within* the most innovative firms in the economy as firms proactively use pay-for-performance to retain the most valuable human assets to cope with the increase in knowledge interdependence.

[32] While "firm pay premium" means how much a firm pays a hypothetical worker with average observable and unobservable characteristics, "firm average earnings" is simply the average of the labor earnings of all employees in a given firm.

Theory

There is a general consensus in the existing literature that employee entrepreneurship is driven by profitable opportunities that are unexploited by the focal firm (Franco and Filson, 2006; Freeman, 1986; Hellmann, 2007; Klepper and Thompson, 2010; Gambardella et al., 2015). From a purely "technological" perspective, given that knowledge inheritance is a major source of "spinout" formation (Agarwal et al., 2004; Golman and Klepper, 2016), potential founders embedded in more integrated knowledge environment might be facing higher costs to replicate or transfer existing know-hows into a new firm. In particular, potential employee entrepreneurs could face greater coordinative challenges to transfer knowledge embodied in a complex system (Grant, 1996; Nickerson and Zenger, 2004). Indeed, this was my initial conjecture as to the reasons why more knowledge interdependent firms may spawn fewer startups.

However, as already noted, if the story were just about overcoming the technology (coordinating knowledge transfer) barriers, we would expect higher-human capital individuals to be more, not less likely to found startups from high-knowledge interdependent parent firms. The empirical results presented in Chapter 2, however, have demonstrated that precisely the opposite is the case.

The novel fact uncovered by the empirical analysis above, that high-performing individuals are less, not more likely to found startups if they work in firms with higher knowledge interdependence, suggests that while individuals with high human capital might be more capable of transferring existing knowledge or/and recreating complementary assets in new firms (Agarwal et al., 2004), such individuals are also more likely to possess indispensable human assets that are complementary to the firm's assets for value creation. Moreover, such firms relying on high knowledge interdependence might give strong incentives to those individuals to prevent them from leaving, presumably because it would be more detrimental for parent firms to lose such individuals.

The above argument suggests that if the internal labor market is efficient, earning rankings (residuals) could be viewed as a proxy not only for human capital in general but more specifically for the individual's capability of mobilizing complementary assets and resources. The fact that higher knowledge interdependence lowers the relative likelihood of startup formation by individuals with high earning rankings (residuals) by more than for those with lower earning rankings (residuals) thus points toward the influence of complementarity between high-performing individuals and the firm's assets. Building on these insights, I present below a simple theoretical construction where higher knowledge interdependence increases the dependence of the firm's assets on high-performing individuals' core knowledge, increasing the bargaining power of those individuals.

A Non-mathematical Overview of the Model

This section verbally describes the formal model and propositions 1-3 below to explicate the underlying logic and intuition. Firms compare the marginal costs against the marginal benefits when setting a compensation scheme for their employees, to influence their decisions of staying or leaving the firm to form a startup.[33] On the one hand, the marginal benefits brought about by a higher wage rate come from the increase in expected marginal productivity as employees are more likely to stay. Furthermore, these incremental marginal benefits would be higher if the focal employees have higher capability and productivity, and they are ranked higher on the job ladder. On the other hand, the marginal cost is just the expected cost of labor that firms need to pay for stayers. Then, in equilibrium, the firm sets an optimal compensation scheme to balance marginal benefits vs. marginal costs such that the expected profit is maximized.

The production function in the model starts from the standard setup where there are increasing returns to managerial ability (Lucas, 1978). Building on the findings that knowledge integration improves firm performance by deterring imitation as complex knowledge is difficult and costly to transfer (Toh and Ahuja, 2022; Rivkin, 2000; Rivkin and Siggelkow, 2002; Grant, 1996; Nickerson and Zenger, 2004), I modify this setup by introducing a crucial assumption that such increasing returns are proportionate to the fraction of the output produced with interdependent knowledge. This means that capable managers (empirically proxied by high-performing, high-ranking employees) allow firms with higher knowledge interdependence to create and appropriate higher value on the product market. In its turn, this means that the firm will benefit by raising the wage, especially for high-performing, high-ranked employees on whom it relies for increasing returns in its knowledge-interdependent production if a higher wage prevents those employees from leaving and taking away important complementary

[33] I abstract from other types of mobility, such as moving to an established competitor to concentrate on the main story.

assets with them (Campbell et al., 2012). Thus, in equilibrium, the marginal benefits of setting a higher wage are higher in firms with higher knowledge interdependence, and even more so for the highest-level talent.

Hence, the outcome that knowledge complexity is indeed a key factor behind the lower rate of employee entrepreneurship arises not simply because of higher barriers to new businesses. It occurs because the profit-maximizing equilibrium in internal labor markets pushes incumbent firms to implement incentive schemes that keep their most valuable human assets from leaving to start their own ventures. I now present an economic model to explicate this mechanism.

Model Setup

Consider an incumbent firm that hires knowledge workers and managers to produce innovative products in a competitive market. Without loss of generality, the market price of one unit of the final output is normalized to unity. Since the focus of this paper is on the interdependence of knowledge as an important input factor of the production process, I assume the production process consists of two types of tasks. One can be accomplished by workers themselves without any coordination by the managers; the other needs knowledge inputs from both workers and managers, which are interdependent to the degree of k Thus, the parameter $k \in [0,1]$ can be interpreted as the fraction of time during which interdependent problem-solving involving collaboration and communication between workers and managers is needed. The role of managers is thus to coordinate interdependent knowledge production in the firm. Indeed, this is in the spirit of the classic model of efficient knowledge hierarchy, where knowledge production requires managers to specialize in exceptions and communicate specialized knowledge to production workers as needed (Garicano and Rossi-Hansberg, 2015, 2004; Garicano, 2000). Assume the productivity of autonomous production by knowledge workers is $y_0(n)$, where n is labor input of knowledge workers and $y_0'(n) > 0$ and $y_o''(n) > 0$. Meanwhile, the productivity of collaborative production involving both workers and managers with human capital θ is $y(n, \theta)$, which is assumed to be monotone increasing and concave in θ. The span of control of managers is fixed so that the number of managers in the firm is normalized to unity. Conditional on all employees staying in the firm, the total output is given by $(1-k)y_0(n) + ky(n, \theta)$.

For simplicity, assume the focal firm is endowed with a fixed number of knowledge workers n with an exogenously determined wage rate $w_0(n)$. Managers, however, are assumed to have human capital and knowledge that are firm-specific and not transferable to other incumbent firms. Thus, the focal firm has monopsony power over managerial human assets, and managers can choose either to work for the incumbent firm with a wage offer w_θ or leave it to start their own business ventures, i.e., form a startup.

I assume that the value of a startup to the manager, $\tilde{\lambda} = \theta - ck + \tilde{v}$ is comprised of three components -- the manager's ability (human capital) θ[34], the cost of creating a startup ck, and the quality of the business idea $\tilde{v}$. The assumption that the value of startups is an increasing function of managerial ability (human capital) reflects

[34] While θ is assumed to be non-transferable to other incumbent firms, I assume, in the spirit of Golman and Klepper (2016) that the manager can utilize it in his/her own business venture.

received wisdom that high-ability employees are better at mobilizing and recreating complementary assets to start new business ventures (Campbell et al., 2012; Carnahan et al., 2012; Agarwal et al., 2016). The expected value of the manager's outside option is also assumed to be negatively associated with the degree of knowledge interdependence in the incumbent firm $c > 0$.[35] This is because it is costly and difficult to transfer complex knowledge to a new organization (Ganco, 2013; Rivkin, 2000; Rivkin and Siggelkow, 2002; Grant, 1996; Nickerson and Zenger, 2004). Thus, c can be thought of as the marginal cost, i.e., additional labor market frictions for potential founders to transfer or recreate complementary assets and capabilities needed to form a startup as knowledge interdependence rises.[36]

The quality of the business idea $\tilde{v}$ is assumed to be private information of the manager while the incumbent (potential parent) firm only knows the distribution $F(\tilde{v})$. Needless to say, the assumption that that manager knows the true value of $\tilde{v}$ is not crucial as all the qualitative results would hold as long as the manager has relatively better information about $\tilde{v}$ than his/her parent firm. To simplify the exposition of the model and focus on the main mechanism, I assume that $\tilde{v}$ is uniformly distributed on

[35] Indeed, $c > 0$ is not necessary to derive the negative relationship between knowledge interdependence and employee entrepreneurship. A sufficient condition, as will be shown later, is that $y(n, \theta) + c > 0$. In other words, it is possible for $c < 0$, i.e., knowledge interdependence lowers the cost of entrepreneurial entry, as long as the marginal productivity of the firm on the product market is sufficiently high such that in equilibrium, the incremental wage would still be higher than the incremental value of startups.

[36] A more general way to model the cost structure of founding a startup is to let $\tilde{\lambda} = c(\theta, \mathrm{k}) + \tilde{v}$, where $c(\theta, \mathrm{k})$ is increasing in k but decreasing in θ. Moreover, $\frac{d^2 c(\theta,k)}{d\theta dk} < 0$ such that the higher costs led by higher k is mitigated by higher capability θ. This alternative assumption would indeed reinforce Proposition 1 below. More details can be provided upon request.

$[-\sigma, \sigma]$.[37] Note that because the domain of the assumed probability distribution is bounded by $-\sigma$ from below and by σ from above, if a firm offers a wage rate below $\theta - ck - \sigma$, all managers with ability θ will "spinout" with probability one; if the wage offer is above $\theta - ck - \sigma$, all managers with ability θ will stay as waged employees with probability one.

The firm then makes a take-it-or-leave-it wage offer to the manager. The manager accepts it if $w_\theta \geq \tilde{\lambda}$ or becomes an entrepreneur otherwise. It is easy to see then that the probability of transition into entrepreneurship is given by $1 - F(w_\theta - \theta + ck)$ (the manager leaves if their privately observed realization of $\tilde{\nu}$ is higher than $w_\theta - \theta + ck$; otherwise, they stay). Moreover, the firm's expected output from collaborative production is $ky(\theta, n)$ if the manager stays, and zero if the manager turns down the wage offer. The unconditional expected output from collaborative production is then given by $ky(\theta, n)F(w_\theta - \theta + ck)$. Thus, the incumbent firm faces the following problem to maximize expected profit π by choosing an optimal wage rate w_θ^*:

$$\max_{w_\theta} \pi = (1-k)y_0(n) + ky(n, \theta)F(w_\theta - \theta + ck) - w_0(n)$$

$$- F(w_\theta - \theta + ck)w_\theta$$

$$s.t. \;\; n \geq 0, \;\; \theta - ck - \sigma < w_\theta < \theta - ck + \sigma$$

[37] The uniform distribution is chosen to simplify the exhibition of higher-order partial derivatives in the model and to explicate key ideas with the most concise format. The theoretical implications of the model do not rely on the functional form of the probability distribution function. All the results hold as long as $F(\tilde{\nu})$is log-concave. Indeed, most of the widely used probability distributions in economics and social science research are log-concave. For example, Uniform, Normal, Exponential, Logistic, Extreme Value, Laplace, Power Function, Weibull, Gamma, Chi-Squared ($c \geq 2$), Chi ($c \geq 1$), Beta ($\nu \geq 1, \omega \geq 1$), Maxwell, Rayleigh, Pareto, and Lognormal distributions all have log-concave CDFs (Bagnoli and Bergstrom, 2006).

In essence, the firm is facing a trade-off between losing the manager and paying higher wages in order to maximize its expected profit. To rule out potential corner solutions, I further assume that $\theta - ck > \sigma$ and $\theta - \sigma < k[y(n,\theta) + c] < \theta + 3\sigma, \forall k, n, \theta, \sigma, y(\cdot)$ such that there exists a unique equilibrium wage $\mathrm{w}_\theta^* \in (\theta - ck - \sigma, \theta - ck + \sigma)$ and $\mathrm{w}_\theta^* \in R^+$.

Model Analysis

As mentioned in the non-mathematical model description, the firm's problem is to balance the expected marginal benefits and marginal costs by setting an optimal wage given the managers' capability θ, the span of control n, and the degree of knowledge interdependence k. In other words, the equilibrium wage w_θ^* is determined by the following first-order condition of the maximization problem:

$$\frac{\partial \pi(\mathrm{w}_\theta^*)}{\partial\, w_\theta^*} = \mathrm{ky(n,\theta)f(w_\theta^* + ck - \theta)} - \mathrm{w}_\theta^* f(\mathrm{w}_\theta^* + \mathrm{ck} - \theta) - \mathrm{F(w_\theta^* + ck - \theta)}$$

$$= 0 \Longrightarrow ky(n,\theta) f(w_\theta^* + ck - \theta)$$

$$= F(w_\theta^* + ck - \theta) + f(w_\theta^* + ck - \theta) w_\theta^*$$

where $ky(n,\theta) f(w_\theta^* + ck - \theta)$ and $F(w_\theta^* + ck - \theta) + f(w_\theta^* + ck - \theta) w_\theta^*$ are the marginal benefits and marginal costs, respectively, for the firm if it raises the wage w_θ^* of managers. The left-hand side is just the expected extra revenue the firm receives from a marginal increase in the probability of the manager staying with the firm, while the right-hand side is the total expected marginal cost of a higher wage.

Under the assumption that $\tilde{v}$ is uniformly distributed on $[-\sigma, \sigma]$, the equilibrium wage offered to a manager with capability θ can be written as:

$$w_\theta^* = \frac{1}{2}(ky(n,\theta) + \theta - ck - \sigma)$$

Moreover, the equilibrium profit of the incumbent firm can be written as:

$$\pi^* = \pi_0 + \frac{1}{8\sigma}(ky(\theta,n) - \theta + ck + \sigma)^2$$

where $\pi_0 \equiv (1-k)y_0(n) - w_0(n)$ is the firm's profit from the independent production of knowledge workers. Note that the proportion of the firm's profit coming from collaborative production is monotonically increasing in k from 0 to 2σ conditional on the existence of an interior solution, i.e., $\theta - \sigma < k[y(n,\theta) + c] < \theta + 3\sigma$.

Now, we turn to our main interests of the relationship among knowledge interdependence k, the probability of entrepreneurial entry $\rho \equiv 1 - F(w_\theta^* + ck - \theta)$ and their average quality, and the equilibrium wage w_θ^*, conditional on different levels of managerial capability θ. To begin with, the following proposition shows the relationship between knowledge interdependence and the equilibrium wage.

Proposition 1. *In equilibrium, firms with higher knowledge interdependence k pay higher wages w_θ^* to their managers if and only if $y(n,\theta) > c$. Moreover, the increase in the equilibrium wage rate is increasing for managers with higher θ. Formally:*

$$\frac{dw_\theta^*}{dk} > 0 \Leftrightarrow y(n,\theta) > c,\ \frac{d^2 w_\theta^*}{dk d\theta} > 0$$

Proof. Since $\tilde{v} \sim \mathrm{U}[-\sigma,\sigma]$, equilibrium wage can be derived from the first-order condition with respect to w_θ:

$$\frac{\partial \pi}{\partial w_\theta} = ky(n,\theta)f(w_\theta^* + ck - \theta) - w_\theta^* f(w_\theta^* + ck - \theta) - F(w_\theta^* + ck - \theta) = 0$$

$$\Rightarrow w_\theta^* = ky(n,\theta) - \frac{F(w_\theta^* + ck - \theta)}{f\left(w_\theta^* + ck - \theta\right)}$$

$$\Rightarrow w_\theta^* = \frac{1}{2}(ky(n,\theta) + \theta - ck - \sigma)$$

Totally differentiating the above with respect to k, we obtain:

$$\frac{dw_\theta^*}{dk} = \frac{1}{2}[y(n,\theta) - c] > 0 \Leftrightarrow y(n,\theta) > c$$

Moreover,

$$\frac{d^2 w_\theta^*}{dkd\theta} = \frac{1}{2} y_\theta'(n,\theta) > 0$$

The firm's problem is to provide an optimal wage offer such that the expected marginal benefit equals the expected marginal cost. While the expected marginal benefit, i.e., $ky(n,\theta)f(w_\theta^* + ck - \theta)$, is strictly increasing in knowledge interdependence k, the expected marginal cost is also increasing in k as it will lower the probability of entrepreneurial entry directly by imposing a higher cost of founding a new business, i.e., $c > 0$. Thus, firms are willing to pay a higher wage premium if and only if the marginal output in the product market is greater than the labor market friction caused by a unit increase in knowledge interdependence. Intuitively, while higher knowledge interdependence directly reduces the likelihood of transition into entrepreneurship by imposing a higher cost of entrepreneurship, the incumbent firm is incentivized to further reduce the risk of losing its manager by offering higher wages when the marginal profit of coordinated production, i.e., $y(n,\theta) > c$, is high enough. Furthermore, since $\frac{dw_\theta^*}{dk} = \frac{1}{2}[y(n,\theta) - c]$, the amount of optimal wage increments should be exactly one-half of the productive output of collaborative production net of the labor market friction if the private information of startup quality is uniformly distributed. Thus, firms and managers share the rents from extra output equally in this simple setup.

It is also easy to see that since collaborative production involving high-ability managers is assumed to be more productive, i.e., $y'_\theta(n, \theta) > 0$, the marginal benefits are even higher for managers with higher θ. This is exactly what we saw in the empirical results in the previous sections, that there the within-firm earning dispersion is higher in firms with higher knowledge interdependence. In addition, the equilibrium wage implies a higher wage premium not just to more capable managers, i.e., those with higher θ, but also to those with the larger span of control n. Interestingly, even though the equilibrium wage of managers will be lower as σ increases, i.e., the private information of startup quality becomes noisier, the relationship might actually be reversed for highly productive firms if the account is taken of the fact that with less noise, the wage required to keep the manager in the firm with probability one ("the-corner-solution" wage) is lower than the same wage if the private information is noisier. Intuitively, for firms that are highly productive, i.e., with very high $ky(\theta, n)$, it is always optimal to ensure that the manager will accept the wage offer and stay in the firm with probability one. This means a corner solution will be reached as σ approaches its highest possible value. In such corner solutions, the equilibrium wage is $\theta - ck + \sigma$, which is obviously, increasing in σ.

Next, I show the relationship between knowledge interdependence and the probability of entrepreneurial entry. Since a higher wage incentive is provided and the costs of entry increase in k, the model also implies that the probability of entrepreneurial entry would be lower in firms with higher knowledge interdependence, and the effect should be more pronounced for high-ability managers. Formally, we have the following proposition.

Proposition 2. *In equilibrium, higher knowledge interdependence is associated with lower likelihood of employee entrepreneurship. Moreover, this negative relationship is more pronounced for managers with higher* θ. *Formally, we have:*

$$\frac{d\rho}{dk} < 0, \qquad \frac{d^2\rho}{dkd\theta} < 0,$$

where $\rho = 1 - F(w_\theta^* + ck - \theta)$ is the probability of transition into entrepreneurship.

Proof. This is obvious by taking the first order derivative of ρ which gives:

$$\frac{d\rho}{dk} = -F'(w_\theta^* + ck - \theta)\left(\frac{dw_\theta^*}{dk} + c\right) = -\frac{1}{4\sigma}[y(n,\theta) + c] < 0$$

Also,

$$\frac{d^2\rho}{dkd\theta} = -\frac{1}{4\sigma}y'_\theta(n,\theta) < 0$$

It is straightforward to see that the probability of transition into entrepreneurship is lower as knowledge interdependence increases regardless of whether the equilibrium wage is higher or lower. It is obvious that when equilibrium wage is higher, i.e., $\frac{dw_\theta^*}{dk} > 0$, the probability of transition into entrepreneurship is lower due to higher opportunity costs and higher entry costs. More interestingly, even when c is high enough such that $\frac{dw_\theta^*}{dk} < 0$, the rises of entry costs will still be larger than the reduction of equilibrium wage, which still makes waged jobs relatively more attractive to the manager. More importantly, the theoretical derivation above shows that the equilibrium outcome of a stronger negative effect of higher knowledge interdependence on the likelihood of transition into entrepreneurship is indeed driven by the fact that high-performance, high-ranked individuals, i.e., those with high θ, contribute to the firm's collaborative production process with higher marginal productivity, i.e., $y'_\theta(n, \theta) > 0$. In other words, the formal model internally validates the conjecture that the higher incremental wage for high-ability managers in firms with higher knowledge interdependence is indeed a plausible mechanism for the more pronounced, negative relationship between knowledge interdependence and startups founded by high-performing, high-ranked individuals.

Last but not least, the model also has implications on startup performance that are consistent with empirical findings in the previous Chapter.

Proposition 3. *Conditional on startup formation, the expected quality of the business idea* $\bar{\nu} \equiv \mathrm{E}[\tilde{\nu} | \tilde{\nu} > w^*_\theta + ck - \theta]$ *is positively associated with the degree of knowledge interdependence* k *of the parent firm, and the effect is more pronounced for founders with higher capability* θ*. Formally,*

$$\frac{d\bar{\nu}}{dk} > 0, \qquad \frac{d^2\bar{\nu}}{dkd\theta} > 0.$$

Proof. First, it can be shown that the threshold startup quality ν_c that makes a manager with ability θ indifferent between waged jobs and entrepreneurship is:

$$\nu_c = w^*_\theta + ck - \theta$$

Thus, the average quality of the business idea conditional on the formation of a startup can be written as:

$$\bar{v} \equiv E[\tilde{v}|\tilde{v} > w_{\theta}^{*} + ck - \theta] = \frac{1}{2}(w_{\theta}^{*} + ck - \theta + \sigma)$$

We know from Proposition 1 that:

$$\frac{dw_{\theta}^{*}}{dk} = \frac{1}{2}[y(n, \theta) - c]$$

Thus,

$$\frac{d\bar{v}}{dk} = \frac{1}{2}\left[\frac{dw_{\theta}^{*}}{dk} + c\right] = \frac{1}{4}[y(n, \theta) + c] > 0$$

And,

$$\frac{d^2\bar{v}}{dkd\theta} = \frac{1}{4}y_{\theta}'(n, \theta) > 0$$

The results follow similar intuition as before. Specifically, it is obvious that higher signal of the quality of startup idea ν is needed for the manager to leave the incumbent firm when higher knowledge interdependence raises both the equilibrium wage and the cost of entry. Moreover, even if the output of coordinated production $y(n, \theta)$ is not large enough relatively c to justify a higher equilibrium wage, the manager who was indifferent between accepting the wage offer and forming a startup will be more inclined to stay in the incumbent firm as higher knowledge interdependence leads to even higher entry costs. As a result, the relative higher wage offered in firms with higher knowledge interdependence leads to a pure selection on the quality of startups that are eventually formed. Importantly, this selection effect is stronger for individuals with higher managerial ability.

In sum, the theoretical model suggests the best alternative explanation for all the empirical findings in the previous sections. Furthermore, it also generates a testable prediction that firms will pay wage premium if and only if the output in coordinated production is greater than the marginal cost of the labor market friction generated by higher knowledge interdependence (see the first part of Proposition 1). We can think of the output in coordinated production as driving the firm's demand for its top-ranked employees, while the marginal cost of the labor market friction is shifting the demand for entrepreneurs (and hence, affecting their supply to incumbent firms). Thus, higher knowledge interdependence increases both the firm's demand for high-level human capital up and the supply of high-level human capital in the internal labor market at the same time. Whether the new equilibrium results in higher or lower compensation for top-ranked employees in more knowledge-interdependent firms is thus an empirical question that I leave for future studies.

[Insert Figure 9 around here.]

Figure 9 uses a numerical example to illustrate the main propositions of the model when the impact of knowledge interdependence on labor market friction is negligible, i.e., $c = 0$, and how equilibrium wage and the probability of transition into entrepreneurship are affected by different levels of knowledge interdependence for two different values of managerial ability. As can be seen from Figure 1, as knowledge interdependence increases, the equilibrium wage for managers is strictly increasing but with a steeper slope for high-θ managers (the red dashed line) as compared to low-θ managers (the red solid line). At the same time, the probability of transition into entrepreneurship is decreasing in k faster for high-θ managers than for low-θ managers.

It is also noteworthy that the probability of transition into entrepreneurship for high-θ managers could be higher than for low-θ managers if the firm's knowledge interdependence is sufficiently low as illustrated in Figure 9. This is due to the potential mismatch between managers and firms as θ is an exogenous parameter in the current version of the model. Specifically, highly capable managers in firms having low knowledge interdependence could have better outside option while the firm does not value the manager's capability of coordinating tasks given that the level of knowledge interdependence is low. The manager's human capital, however, becomes extremely valuable for firms having higher knowledge interdependence, which in turn leads to higher wage and lower probability of transition into entrepreneurship as k increases. Also, conditional on startup formation, the average quality of startups founded by previous employees from high-knowledge-interdependence parents will be higher than those emanating from parent firms with lower knowledge interdependence in any event.

Empirical Analysis

To find supporting evidence for the predictions of the model, I utilize information on individual earnings from the LEHD and firm knowledge interdependence measures constructed in Chapter 2 to further investigate the underlying mechanism as to how incumbent firms retain the most valuable human assets within the firm and prevent them from leaving. In what follows, I focus on the relationship between knowledge interdependence and compensations, i.e., predictions from Proposition 1, as Propositions 2 and 3 have already been shown as consistent with empirical findings in the previous chapter.

Within-firm Compensation Dispersion

The negative impact of losing key human capital on firm performance, especially where there are strong complementarities with the firm is well understood (Agarwal et al., 2016). If higher knowledge interdependence creates such stronger complementarities, it calls for incumbent firms who rely on such interdependence to provide stronger incentives to retain key personnel who have control of and access to important complementary assets. Thus, firms with higher knowledge interdependence are expected to pay more especially to their high-level employees. Moreover, if the proposed mechanism in the theoretical model is plausible, compensation dispersion between high-ability and low-ability workers should be greater in those firms replying on higher knowledge interdependence. I examine this prediction in this section.

Specifically, I estimate the following earnings regressions:

$$Y_{i,j,t+1} = \beta_0 + \beta_1 K_{j,t} \times r_{i,j,t} + \gamma X_{i,t} + \delta_t \times d_j + \mu_{i,j,t}$$

where $Y_{i,j,t+1}$ is the log of real annualized earnings of individual *i* in firm *j*, year *t*; $K_{j,t}$ is the knowledge interdependence of firm *j* in year *t*; $r_{i,j,t}$ is the earning rankings of individual *i* in firm *j*, year *t*; $X_{i,t}$ is a set of individual controls including earning rankings, as well as age and its quadratic term, current job tenure and its quadratic term, gender, and race; and $\delta_t \times d_j$ is the firm-year fixed effect. Robust standard errors are clustered at the firm level.

[Insert Table 15 around here.]

Table 15 shows the results. Column (1) shows that the coefficient on the interaction term between firm knowledge interdependence and individual earning rankings is positive and significant. This means that earnings are increasing more with relative rankings in firms with higher knowledge interdependence. Simply put, top-ranked individuals are earning more in firms with higher knowledge interdependence, all other things equal. Column (2) shows the earning gaps between employees ranked in top 5% within the firm vs. other employees. The coefficient estimates imply that one standard deviation increase in knowledge interdependence increases the gap between the top 5% earners and the rest of the firm's employees by 2.4%. Interestingly, as shown in Column (3), those ranked in the bottom ten percent do not seem to be significantly affected by the degree of knowledge interdependence. If anything, they seem to be paid even less in firms with higher knowledge interdependence.

Individual-level analysis suggests that firms with higher knowledge interdependence might be using greater compensation dispersion to discourage high-performing individuals from leaving. To verify this result, I constructed the Gini coefficient, which is widely used in studies of compensation dispersion and in labor economics (Carnahan et al., 2012; Bloom, 1999; Bloom and Michel, 2002; Donaldson and Weymark, 1980; Shaw et al., 2002). The Gini coefficient ranges from 0 (absolute equality) to 1 (absolute inequality), and it measures half the relative mean difference of the pay of any two randomly selected employees from a firm's pay distribution. It is calculated as:

$$G = \frac{2\sum_{i=1}^{n} i\, y_i}{n\sum_{i=1}^{n} y_i} - \frac{n+1}{n}$$

where y_i is the annualized real earnings of the *i*th ranked individual in a firm and is indexed in non-decreasing order and *n* is the number of employees in the firm. For example, $i = 1$ indicates the lowest ranked person, I also calculated the 99-to-5 percentile, 90-to-10 percentile, and 75-to-25 percentile ratios in the firm's earning distribution as alternative measures for income inequality. To investigate whether firms with higher knowledge interdependence have wider earning dispersion, I then estimate the following firm-level OLS regression:

$$W_{j,t} = \alpha + \beta K_{j,t} + \lambda Z_{j,t} + \delta_t \times \gamma_{state} \times \nu_{industry} + d_j + \epsilon_{j,t}$$

where $W_{j,t}$ is either the Gini coefficient in firm *j*, year *t*, or the afore-mentioned earnings ratios; $Z_{j,t}$ are firm control variables including the number of technology groups in the firm's patent portfolio, firm size, industry diversification, employment growth, firm age; $\delta_t \times \gamma_{state} \times \nu_{industry}$ is the year-industry-state fixed effect and d_j is the firm fixed effect. Table 16 shows the results.

[Insert Table 16 around here.]

Consistent with findings from the earning analysis at the individual level above, Column (1) shows that the relationship between firm knowledge interdependence and the Gini coefficient is positive and statistically significant. The Gini coefficient at an average firm is around 0.32, and one standard deviation increase in knowledge interdependence would increase it by 0.005, around 2% increase at the mean. However, when the firm fixed effect is included in Column (2), both knowledge interdependence and the number of technology groups become statistically insignificant.

Columns (3) to (8) provide similar insights by replacing the Gini coefficient with earning differentials. Importantly, knowledge interdependence seems to affect within-firm earning dispersion the most at the extreme, i.e., the earning differential between the top 99 and the bottom 5 percentile.

Overall, these results suggest that firms with higher knowledge interdependence have greater within-firm income inequality, while the source of variation is largely coming from cross-sectional differences. There may just be not enough within-firm variation to identify any significant effects from higher knowledge interdependence on earnings schedules when firm fixed effects are included.

Discussion, Open Questions, and Avenues for Future Research

As innovation becomes more complex, requiring interdependent knowledge inputs, firms need to adapt and proactively find efficient ways to manage their human assets. In order to fully grasp the crucial implications of this for economic development and income (re)distribution we need to know how firms cope with the increase in knowledge interdependence. To address this issue, I showed in Chapter 2 that higher knowledge interdependence is by no means working in a simple way as other labor market frictions which merely impose higher cost of entry. Instead, in this Chapter, I show the most plausible theoretical explanation is that higher knowledge interdependence affects both product market competition and firm monopsony power. Importantly, the theory suggests that as long as innovative firms earn enough "entrepreneurial profits" by creating complex combinations (Schumpeter, 1911), they should share this "rent" with their employees, in particular the most talented ones, by paying them higher wage premium. This alternative explanation is consistent with all empirical findings. Namely, higher knowledge interdependence is associated with lower likelihood of employee entrepreneurship, better chance of survival, and higher within-firm income inequality in particular among the most talented entrepreneurs.

This apparent extra layer in the causal chain is important because it leads to a new perspective on declining employee entrepreneurial activities in innovative firms. Specifically, if this decline was solely due to knowledge barriers, as initially conjectured, the implications for innovation activity could indeed be dire. However, if the decline is mediated by higher pay for performance in incumbent firms, a conjecture can be made that increasing knowledge complexity results in simply reallocating the most important innovations from new startups to incumbent firms which are better equipped to handle more complex and interdependent innovation processes. Rising returns to highest-level human capital and the reluctance of individuals possessing such human capital to engage in employee entrepreneurship can then be interpreted as an indicator of such efficient reallocation of innovative activity, while better survival of startups that are still founded by high-human capital individuals who do leave firms with high knowledge interdependence provides further supporting evidence.

Nevertheless, it still remains an open question whether the decline in employee entrepreneurship is associated with an increase in concentration of innovation activities in large incumbent firms, and whether higher retainment of talents leads to better innovation outcomes and firm performance. Specifically, it is imperative to know if it is large firms that are driving the increasing trend of knowledge interdependence by creating inventions relying on new combinations of technological fields that have never been associated with each other. Also, would large firms be more likely to create successful new combinations than startups? I leave these open questions for future research.

Also, the fact that higher knowledge interdependence is associated with higher within-firm compensation dispersion indicates the growing complexity of innovation is an important but understudied aspect of technological change that has been driving up income inequality in the U.S. economy in recent decades (Acemoglu and Autor, 2011; Song et al., 2019). While a much more thorough examination of the impact of knowledge interdependence on income inequality is needed, the current findings provide initial evidence that part of the within-firm income inequality is associated with higher knowledge interdependence and stronger complementarity within highly innovative firms. Importantly, the theoretical framework also predicts that higher knowledge interdependence would also lead to greater between-firm wage inequality if firms earn enough high "entrepreneurial profits," which remains to be fully examined by future research.

This study is not without limitations. First, while the theoretical framing suggests that more complex and interdependent knowledge directly affects entry into entrepreneurship, it may actually lower the barrier to entry in some cases, e.g., e-platforms. This is in particular plausible when incumbent firms deal with the increasing knowledge complexity via modularization since it may be possible to pluck out one module from a complex knowledge system and turn it into an independent and successful business. In other words, the degree of modularity of an incumbent firm's knowledge network could be another important aspect affecting potential entrepreneurs' decisions to start their own business ventures. While this potential effect is absorbed by the firm, industry, and year fixed effects if most of the source of variation of the degree of modularity originates across industries, this may be an important aspect to embrace and investigate in the future research.

Also, one implicit assumption made in the current analysis is that all employee entrepreneurs are driven by technological opportunities, which might not be true in reality even for highly innovative parent firms because entrepreneurs have different motivations to start a new venture. At this stage, data limitations (lack of an inventor bridge between USPTO and LEHD data) prevent me from identifying employee entrepreneurs who are also inventors. Once the USPTO-LEHD bridge becomes available, it will be possible to conduct more detailed analyses, including comparing inventors and non-inventors among employee entrepreneurs emanating from highly innovative parent firms. Note, however, that the current results, which mix employee entrepreneurs who are driven by technological opportunities with others who are not, are more likely to underestimate, not overestimate, the relationship between knowledge interdependence and the likelihood of entrepreneurial entry of inventors as non-inventors might not be affected by the knowledge structure in innovation activities at all. Indeed, the stronger effects on high residual earners are consistent with this reasoning as high-performing founders are more likely to be high-skilled knowledge workers discovering unexploited technological opportunities.

Furthermore, the measurement of firm knowledge interdependence might be affected by changes in the patent examination procedures at the USPTO, e.g., changes in the easiness of search, and the transition from USPC to CPC around 2013. For instance, if there is a tendency to assign more technology classes to a patent as a result of those changes, this might mechanically drive up the measure of knowledge interdependence. Since industry and year fixed effects are included, however, this should not bias any results as long as the effect of institutional changes affect all patents granted by the USPTO in a similar way within each industry. Nevertheless, it is still important to know how much of the rising knowledge complexity needed for innovation is real. One way to answer this question is to compare the trends in the U.S. with other countries, e.g., the European Patent Office (EPO) data, to eliminate the potential impact of institutional changes peculiar to the United States.

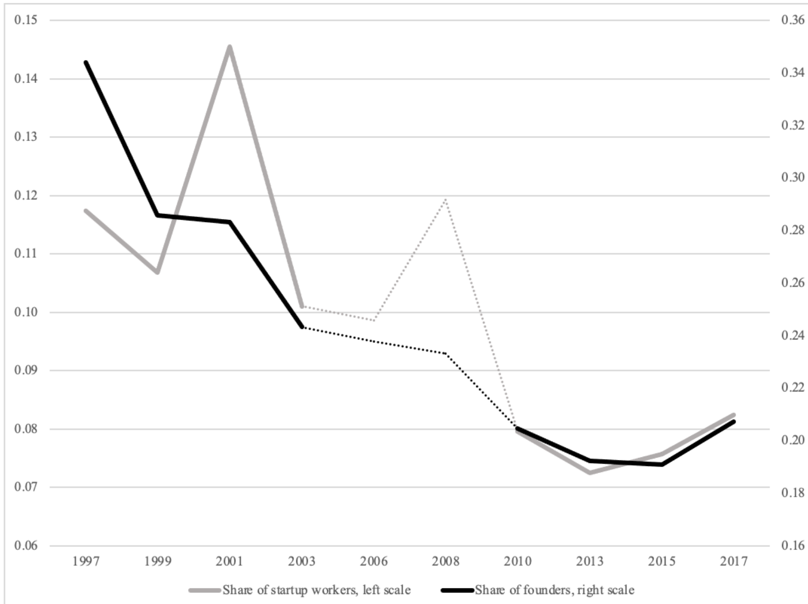
There are several more questions that need to be addressed in future research. First, while the current analysis focuses on the likelihood of transition into entrepreneurship of individual employees, it remains unclear how parent firm knowledge interdependence would affect founding team size. Indeed, a potential entrepreneur does not have to be a "jack of all trades" (Lazear, 2005), but s/he could assemble a team with co-founders having complementary capabilities (Agarwal et al., 2016). Ganco (2013) has found that inventor entrepreneurs with more complex knowledge embodied in them are more likely to pursue team entrepreneurship. Even though the knowledge embodied in individual employees of high knowledge interdependent firms is not necessarily complex, they might still have developed strong complementarity with other workers as greater coordination is needed in such knowledge environment. As high-performing founders could have better access to recruit others (Agarwal et al., 2016), the impact of knowledge interdependence on founding team size, if there is any, could also be moderated by founder characteristics including earning rankings and job tenure, etc.

Indeed, an important question that remains to be addressed with more detailed data on the individual level, e.g., the forthcoming USPTO-LEHD inventor crosswalk, is who are those inventors that found startups? Do they leave their parent firms because they are working on the knowledge components that are peripheral in the parent firm's knowledge network? Or are they more likely to be those who are working on modular components that can be easier to plug out from a complex knowledge network?

Finally, a potentially important line of future examination is the relationship between higher knowledge interdependence and incumbent firm performance. The current results have already provided supporting evidence that firms with higher knowledge interdependence are able to retain high-skilled labor by lowering the likelihood of a "spin-out". While existing studies have investigated the impact of knowledge integration on firm performance and innovative output (Toh and Ahuja, 2022; Ethiraj and Posen, 2014), it remains unclear how strategic human capital management practices moderate or/and mediate this relationship. Indeed, based on the formal theory I developed in this paper, firms with higher knowledge interdependence should have higher profits. Moreover, this relationship should be moderated by the quality of human assets – the coordination capability of managers, or managerial practices that are implemented by the firm. Thus, there are opportunities for future research to investigate if and how highly integrated knowledge structures would influence innovation productivity, labor productivity, firm growth rate, the chances of survival, and so on.

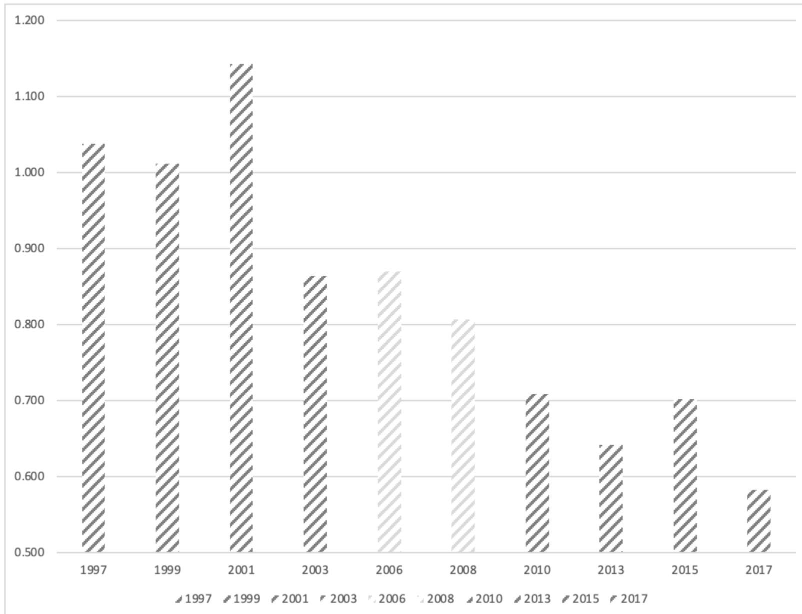
The above is crucial to know to better evaluate the recent declining business dynamism in the U.S. It is possible that as innovation becomes more and more complex, large established firms start to substitute the role of start-ups in pushing forward the technological frontier and driving economic growth. For instance, organizational economists have suggested that the efficient knowledge hierarchy could better deal with complex knowledge needed in the production process (Garicano, 2000; Garicano and Rossi-Hansberg, 2004). From the social planner's perspective, allocating complex innovation activities to large established firms might be efficient especially when most start-ups are inevitably limited in size to achieve the efficient knowledge hierarchy. If this is the case, the declining business dynamism might just be a reflection of technological change and efficient (re)allocation of resources but not necessarily detrimental to technological advancement and economic growth.

Figure 1. Dynamics of the share of PhDs who are founders of startups or employed at startups



Notes: The figure reports the share of PhDs in science and engineering who are founders of startups working full-time with non-zero salaries in the total number of PhDs in science and engineering who are self-employed or owners of established businesses. It also reports the share of PhDs in science and engineering employed full-time with non-zero salaries in new (five years old or less) private for-profit companies compared in the total number of PhDs in science and engineering employed full-time with non-zero salaries in all private for-profit businesses. "Founders" are owners/self-employed at an incorporated startup at the time observed in the data. The business age question was not asked in the 2006 and 2008 surveys. We have imputed the numbers for these years, according to the procedure described in the main text and in the Appendix. The imputed numbers are connected by dotted lines.

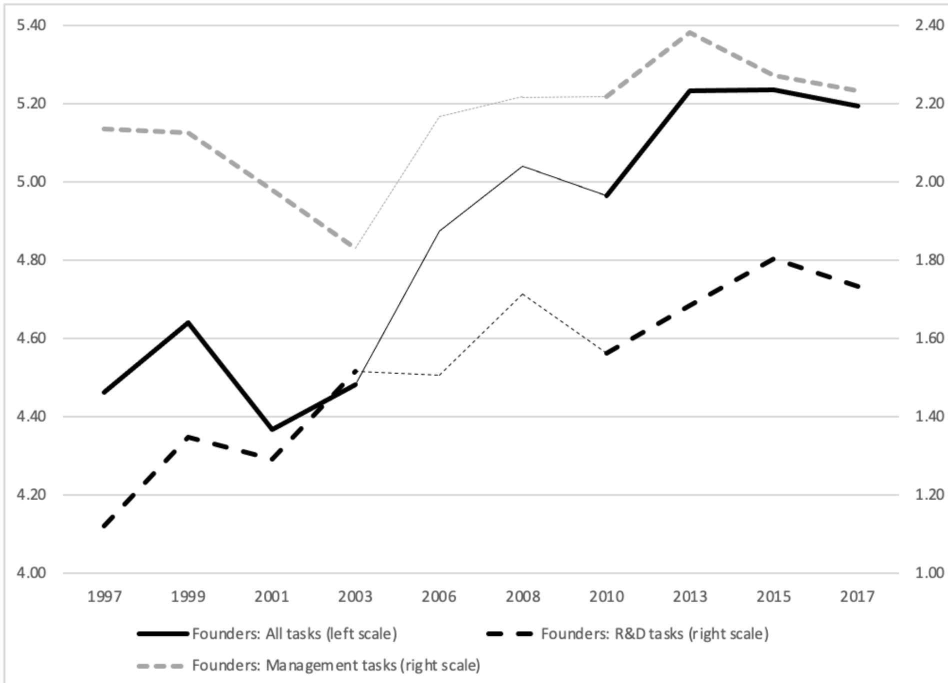
Figure 2. Dynamics of relative earnings of founders with below the median years after PhD



Notes: The figure reports the ratio of earnings of founders with years after PhD at or below the median of 13 years, to earnings of founders with 14 or more years after PhD among PhDs in science and engineering. Lighter shaded bars correspond to imputed salaries for the 2006 and 2008 surveys. See Appendix.

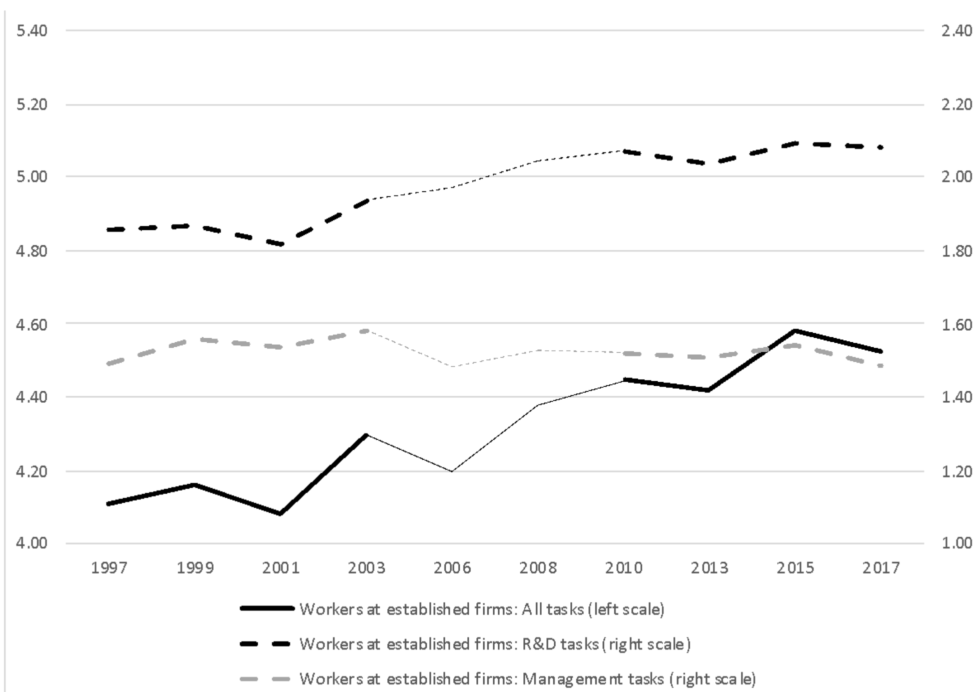
Figure 3

Panel A. The number of all, R&D, and management work activities among founders



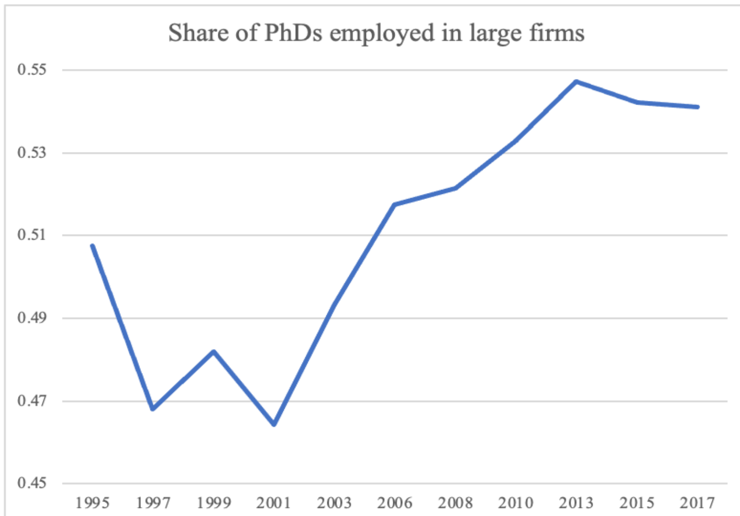
Panel B.

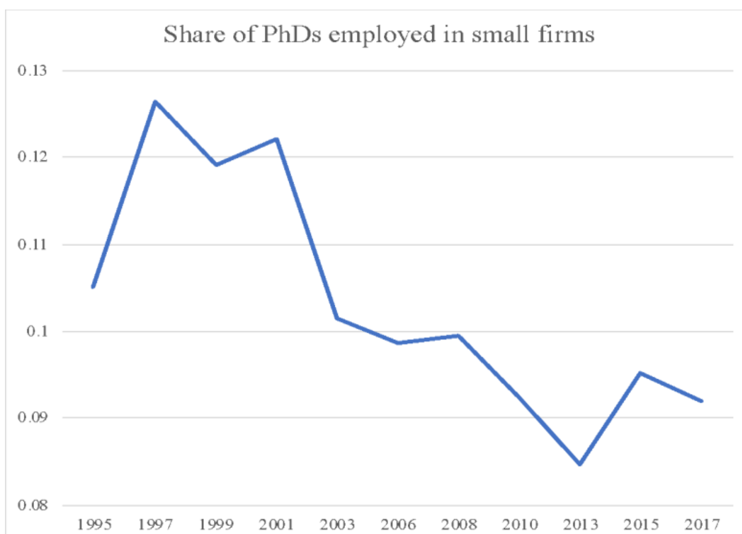
The number of all, R&D, and management work activities among workers at established firms



Notes: R&D activities consist of "basic research," "applied research," "development," and "design." Management activities consist of "accounting, finance, contracts," "human resources," "managing or supervising people or projects," "sales, purchasing, marketing, customer service, public relations," and "quality or productivity management." The total number of work activities includes, in addition to the above, computer programming, employee relations, production, operations and maintenance, teaching as well as other work activities. See Appendix Table A2. Thinner lines connect imputed data for the 2006 and 2008 surveys.

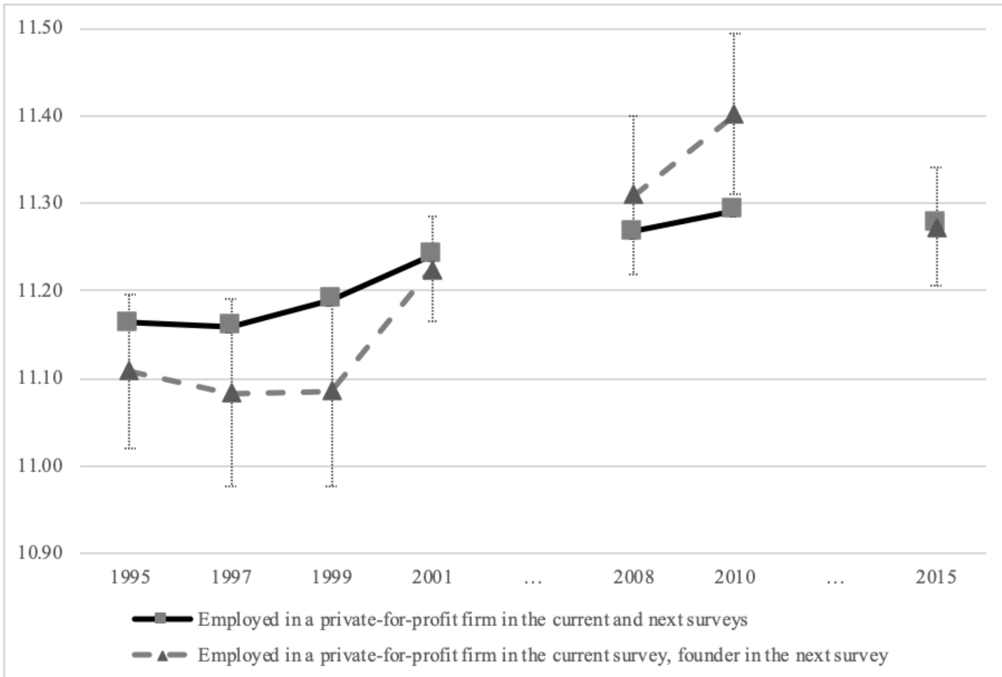
Figure 4. Share of PhDs employed in large and small for-profit firms in total





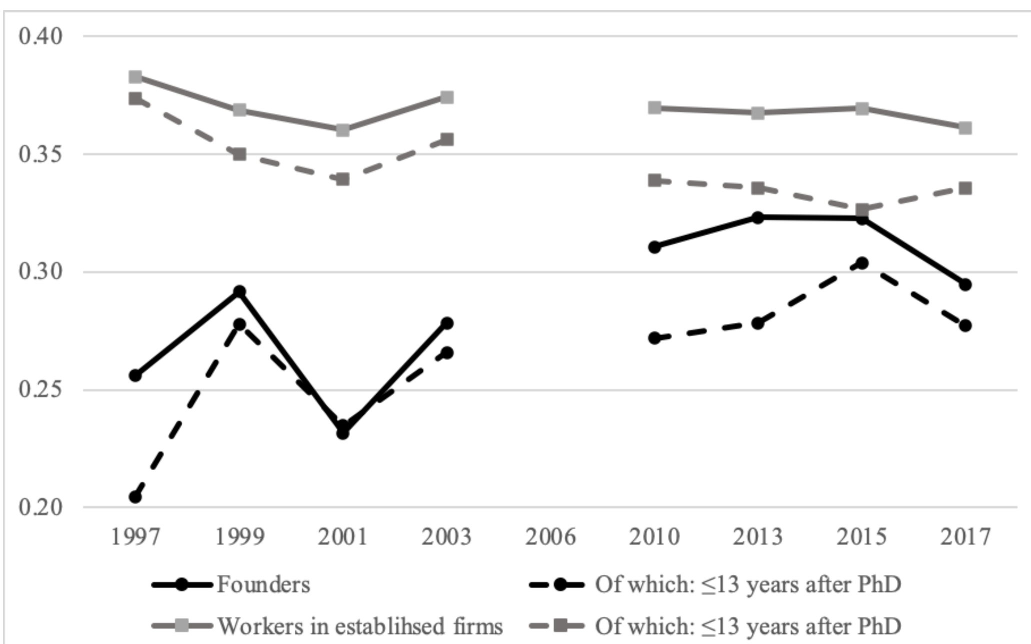
Source: Author calculations using weighted National Science Foundation data. "Large" firms are those with 5,000 or more employees. "Small" firms are those with less than 25 employees, as in Elfenbein, et al. (2010).

Figure 5. Predicted logged pre-choice salaries for those becoming startup founders in the next survey and those remaining as wage workers



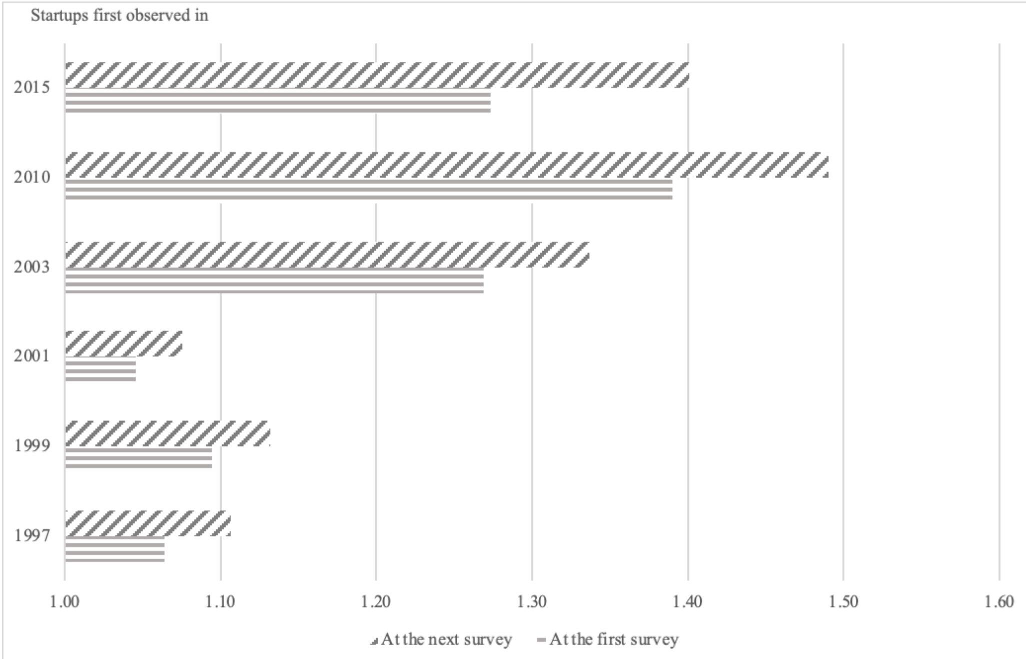
Source: Author estimates based on National Science Foundation data. The bars represent 95 percent confidence intervals.

Figure 6. Share of graduates from top-ranked programs



Source: Author calculations using weighted National Science Foundation data.

Figure 7. Average size of startups in panel data: years of first and next survey



Source: Author calculations using weighted National Science Foundation data.

Figure 8. Differences in fractions for whom salary is very important and for whom independence is very important

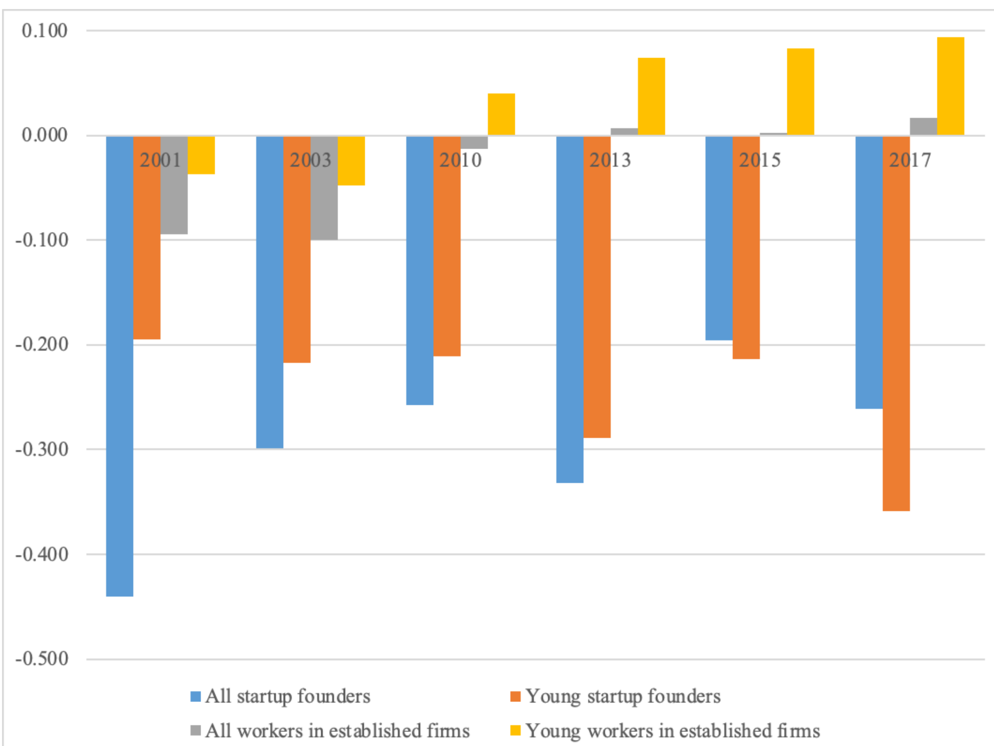
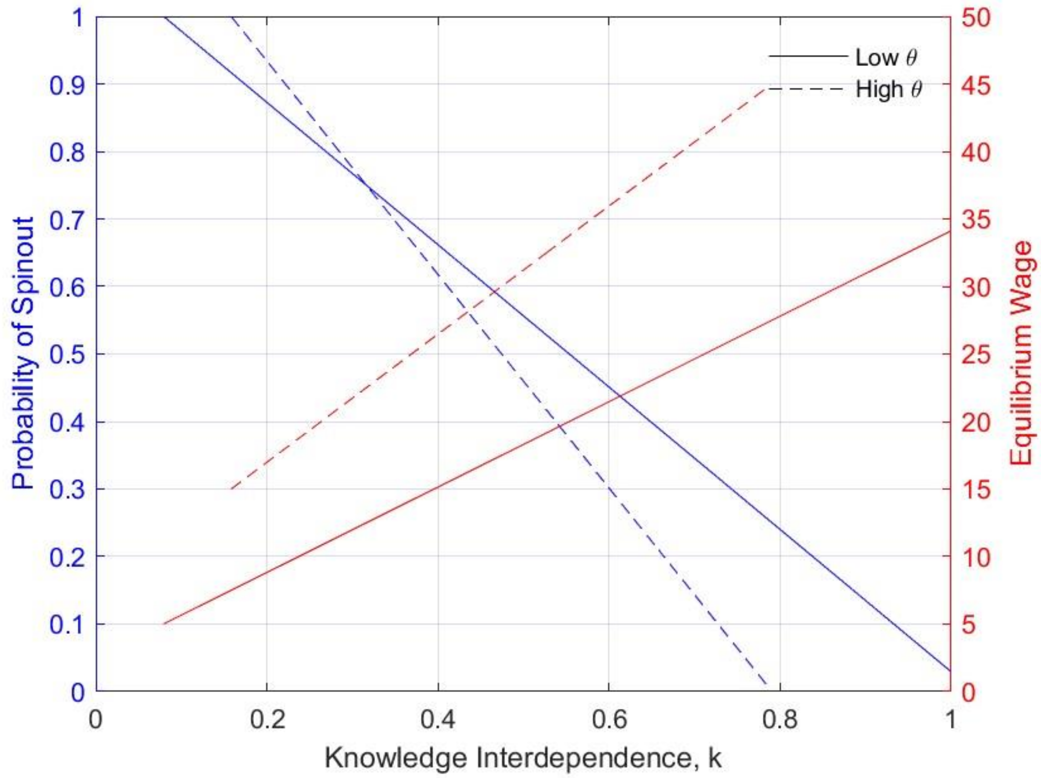


Figure 9: A Numeric Example: Equilibrium Wage and the Probability of Transition into Entrepreneurship



Notes: This figure plots the numeric solutions of the model based on the following sets of parameters. $y(\theta, n) = \theta n^{0.5}$; $\tilde{v} \sim U[15,15]$; $\theta_h = 30$; $\theta_l = 20$; $n = 10$; $\sigma = 15$; $c = 0$ The blue solid line and the blue dashed line are measured on the left y-axis, which indicates the probability of entrepreneurial entry for managers having θ_l and θ_h respectively; The red solid line and the red dashed line are measured on the right y-axis, which indicates the equilibrium wage of managers having θ_l and θ_h respectively.

Table 1.

Panel A. Factors that affect the number of R&D tasks.

	DV: IHS-transformed number of R&D tasks	
	Founders (1)	Workers at established firms (2)
Time trend	0.0109 (0.002)	0.0075 (0.000)
Experience (years after PhD)	-0.0021 (0.002)	-0.0052 (0.000)
Constant	0.905 (0.196)	0.953 (0.048)
Other controls: gender, ethnicity, occupation, U.S. state of employment	Included	Included
Observations	2,101	65,485
R-squared	0.345	0.274

Panel B. Factors that affect the span of control and depth of the hierarchy

	DV: span of control (IHS-transformed # of individuals directly supervised)		DV: depth of hierarchy (IHS-transformed # of individuals indirectly supervised)	
	Founders (3)	Workers at established firms (4)	Founders (5)	Workers at established firms (6)
Time trend	0.0001 (0.004)	-0.0046 (0.001)	-0.0027 (0.004)	0.0032 (0.001)
Experience (years after PhD)	-0.0010 (0.002)	0.0064 (0.001)	0.0051 (0.003)	0.0101 (0.001)
Constant	1.058 (0.405)	0.686 (0.081)	0.717 (0.413)	-0.054 (0.099)
Other controls: gender, ethnicity, occupation, U.S. state of employment	Included	Included	Included	Included
Observations	2,101	65,485	2,101	65,485
R-squared	0.208	0.218	0.264	0.470

Estimation method: OLS using National Science Foundation weights. Robust standard errors in parentheses. Observations with zeroes included using inverse hyperbolic sine (IHS) transformations, defined as $y = ln\left(x + \sqrt{x^2 + 1}\right)$.

Table 2.

Panel A. Factors affecting the number of R&D tasks for employees at large firms.

	DV: IHS-transformed number of R&D tasks	
	(1) Workers at large established firms	(2) Workers at all large firms
Time trend	0.0057 (0.001)	0.0045 (0.000)
Experience (years after PhD)	-0.0048 (0.000)	-0.0051 (0.000)
Constant	1.0095 (0.073)	1.0372 (0.063)
Other controls: gender, ethnicity, occupation, U.S. state of employment	Included	Included
Observations	36,115	47,987
R-squared	0.226	0.222

Panel B. Factors affecting the span of control and depth of the hierarchy.

	DV: span of control (IHS-transformed # of individuals directly supervised)		DV: depth of hierarchy (IHS-transformed # of individuals indirectly supervised)	
	(3) Workers at large established firms	(4) Workers at all large firms	(5) Workers at large established firms	(6) Workers at all large firms
Time trend	-0.0036 (0.001)	-0.0028 (0.001)	0.0065 (0.001)	0.0065 (0.001)
Experience (years after PhD)	0.0077 (0.001)	0.0063 (0.001)	0.0117 (0.001)	0.0099 (0.001)
Constant	0.6277 (0.130)	0.6230 (0.113)	-0.1227 (0.158)	-0.0877 (0.131)
Other controls: gender, ethnicity, occupation, U.S. state of employment	Included	Included	Included	Included
Observations	36,115	47,987	36,115	47,987
R-squared	0.210	0.222	0.489	0.502

Estimation method: Pooled OLS using weighted National Science Foundation data. Robust standard errors in parentheses. Since there are observations with zero number of R&D-related work activities as well as zero span of control and depth of hierarchy, we use the inverse hyperbolic sine (IHS) transformations, defined as $y = ln\left(x + \sqrt{x^2 + 1}\right)$.

Table 3. Earnings of founders and workers at established firms

DV: Log(real earnings)		
	Founders (1)	Workers at established firms (2)
Time trend	-0.0163 (0.006)	0.0039 (0.001)
Experience (years after PhD)	0.0121 (0.009)	0.0277 (0.001)
Experience squared	-0.0005 (0.000)	-0.0005 (0.000)
Experience x time trend	0.0011 (0.000)	0.0001 (0.000)
IHS-transformed number of R&D tasks	0.0488 (0.046)	0.0437 (0.005)
Span of control (IHS-transformed # of individuals directly supervised)	0.0498 (0.029)	0.0466 (0.002)
Depth of hierarchy (IHS-transformed # of individuals indirectly supervised)	0.0820 (0.024)	0.0461 (0.003)
Constant	10.0238 (0.455)	10.5638 (0.042)
Other controls: gender, ethnicity, occupation, U.S. state of employment	Included	Included
Observations	2,101	65,485
R-squared	0.138	0.221

Estimation method: OLS using National Science Foundation weights. Robust standard errors in parentheses. Observations with zeroes included using inverse hyperbolic sine (IHS) transformations.

Table 4. The dynamics of shares of R&D-focused founders and workers in total

Year	(1) Founders	(2) Owners of established incorporated businesses	(3) Workers at established firms	(4) Workers at all large firms
1995				0.600
1997	0.261	0.118	0.509	0.586
1999	0.288	0.150	0.520	0.598
2001	0.316	0.157	0.508	0.578
2003	0.359	0.297	0.512	0.555
2006				0.545
2008				0.566
2010	0.309	0.279	0.524	0.561
2013	0.308	0.285	0.503	0.531
2015	0.373	0.287	0.512	0.553
2017	0.316	0.328	0.517	0.547

Source: Author calculations using weighted National Science Foundation data.

Table 5. Earnings of R&D-focused founders and workers at established firms

DV: Log (real earnings)		
	(1) Founders	(2) Workers at established firms
Time trend	-0.0445 (0.012)	0.0027 (0.001)
Experience (years after PhD)	-0.0161 (0.013)	0.0293 (0.001)
Experience squared	-0.0003 (0.000)	-0.0005 (0.000)
Experience x time trend	0.0026 (0.001)	0.0001 (0.000)
IHS-transformed number of R&D tasks	-0.1521 (0.123)	0.0326 (0.009)
Span of control (IHS-transformed # of individuals directly supervised)	0.1109 (0.054)	0.0372 (0.003)
Depth of hierarchy (IHS-transformed # of individuals indirectly supervised)	0.0549 (0.044)	0.0364 (0.004)
Constant	11.9358 (0.477)	10.7582 (0.048)
Other controls: gender, ethnicity, occupation, U.S. state of employment	Included	Included
Observations	690	33,710
R-squared	0.234	0.179

Estimation method: OLS using weighted National Science Foundation data. Robust standard errors in parentheses. Since there are observations with zero numbers of R&D-related work activities as well as zero span of control and depth of hierarchy, we use the inverse hyperbolic sine (IHS) transformations, defined as $y = ln\left(x + \sqrt{x^2 + 1}\right)$.

Table 6. Growth rates of startups sizes in panel data

DV:	Logged size category in the next survey, minus logged size category in the first observed survey
Time trend	0.016 (0.008)
Logged size category in the first observed survey	-0.917 (0.043)
Constant	0.789 (0.114)
Individual fixed effects	Included
Observations	542
Within R-squared	0.180

Estimation method: Panel estimation with individual fixed effects, using weighted National Science Foundation data (weights for each individual averaged across two surveys). Robust standard errors in parentheses.

Table 7: Summary Statistics - Firm Characteristics

	Mean	S.D.
Knowledge Interdependence	1.733	1.198
Number of Technology Groups	6.77	10.07
Log(Knowledge Interdependence)	1.198	0.5214
Log(Number of Technology Groups)	1.51	0.7713
Log(Firm Employment)	5.066	2.104
Firm Age	20.42	11.26
Young Firm (age ≤ 5)	0.1514	0.3584
Diversification	0.1652	0.249
DHS Employment Growth	0.1213	0.4703
Startup	0.3099	0.4625
Number of Startup	1.03	3.922
Gini Index	0.3138	0.1072
99/5 Earning Differential	2.345	0.7836
90/10 Earning Differential	1.376	0.4663
75/25 Earning Differential	0.7013	0.2807
Rounded N	59,000	59,000

Table 8: Summary Statistics - Individual Characteristics

	All Employees		Primary Founders	
	Mean	S.D.	Mean	S.D.
Spinout founder in year $t+1$	0.001	0.033	-	-
Age	40.8	11.08	37.52	10.83
Job Tenure	4.142	3.198	3.073	2.445
Within-Firm Earning Ranking	50.03	28.86	53.35	33.32
Top 5% Earner	0.050	0.219	0.129	0.335
Earning Residual	0.00	0.53	0.093	0.797
Log(Annualized Earnings)	10.84	0.8067	-	-
Industry Experience	-	-	1.869	1.12
Rounded N	58,860,000	58,860,000	55,500	55,500

Table 9: Time Trend of Knowledge Interdependence

DV: Knowledge Interdependence	(1)	(2)	(3)	(4)	(5)
Time Trend	0.027	0.021	0.026	0.026	0.017
	(0.001)	(0.003)	(0.001)	(0.001)	(0.004)
Log(Firm Size)		-0.019			-0.017
		(0.005)			(0.006)
Time Trend × Log(Firm Size)		0.001			0.002
		(0.0006)			(0.0007)
Young Firm (age ≤ 5)			0.06		0.027
			(0.025)		(0.027)
Time Trend × Young Firm			0.009		0.013
			(0.003)		(0.004)
Diversification				-0.054	0.034
				(0.037)	(0.045)
Time Trend × Diversification				0.005	<0.001
				(0.005)	(0.006)
Constant	1.56	1.654	1.545	1.569	1.637
	(0.009)	(0.025)	(0.01)	(0.01)	(0.03)
Industry & State FE	YES	YES	YES	YES	YES
Rounded Observations	59,000	59,000	59,000	59,000	59,000
Adjusted R-squared	0.065	0.066	0.067	0.065	0.067
Within R-squared	0.008	0.008	0.009	0.008	0.009

Notes: Standard errors in parentheses. Regressions are estimated by using OLS. Industry is classified at the 4-digit NAICS level. The sample consists of all patenting firms with at least two patents over the years 2000-2013 and at least two CPC classes to calculate knowledge interdependence.

Table 10: Time Trend of Employee Entrepreneurship

DV: Startup Formation in Year *t+1*	(1)	(2)	(3)	(4)	(5)
Time Trend	-0.012	-0.011	-0.033	-0.036	-0.043
	(0.001)	(0.001)	(0.004)	(0.002)	(0.002)
Earning Ranking			0.008		
			(0.001)		
Time Trend × Earning Ranking			-0.0001		
			(0.0001)		
Top 5% Earners				2.81	
				(0.134)	
Time Trend × Top 5% Earners				-0.113	
				(0.013)	
Earning Residuals					0.638
					(0.06)
Time Trend × Earning Residuals					-0.022
					(0.006)
Constant	0.386	0.378	1.172	1.126	0.675
	(0.005)	(0.004)	(0.125)	(0.121)	(0.135)
Industry × State FE	YES	NO	NO	NO	NO
Firm FE	NO	YES	NO	NO	NO
Firm × Year FE	NO	NO	YES	YES	YES
Individual Controls	NO	NO	YES	YES	YES
Rounded Observations	59,000	59,000	58,860,000	58,860,000	58,860,000
Adjusted R-squared	0.133	0.359	0.001	0.002	0.001
Within R-squared	0.011	0.01	0.0002	0.0001	0.0003

Notes: Robust standard errors in parentheses. Regressions are linear probability models estimated by using OLS. Industry is classified at the 4-digit NAICS level. The sample consists of all patenting firms with at least two patents over the sample period and having at least two CPC classes such that knowledge interdependence is well-defined. Individual controls include age, age squared, job tenure and its squared term, gender, and race. For exhibition purposes, coefficient estimates and standard errors in Columns 3-5 have been multiplied by 10^3.

Table 11: Knowledge Interdependence and the Propensity of Startup Formation

DV: Startup Formation in Year *t+1*	(1)	(2)	(3)	(4)	(5)
	LP	LP	LP	Logit	Probit
Log(Knowledge Interdependence)	-0.043	-0.035	-0.011	-0.158	-0.104
	(0.005)	(0.005)	(0.006)	(0.025)	(0.015)
Log(Number of Technology Groups)	0.059		0.016	0.176	0.122
	(0.004)		(0.006)	(0.019)	(0.011)
Log(Firm Size)	0.098	0.098	0.082	0.765	0.425
	(0.002)	(0.002)	(0.004)	(0.01)	(0.005)
Diversification	0.165	0.163	0.074	0.496	0.322
	(0.012)	(0.012)	(0.023)	(0.05)	(0.03)
Employment Growth	-0.013	-0.013	-0.021	-0.163	-0.086
	(0.004)	(0.004)	(0.004)	(0.03)	(0.016)
Firm Age	-0.002	-0.002		-0.012	-0.007
	(0.0003)	(0.0003)		(0.001)	(0.001)
Constant	-0.21	-0.13	-0.128	-4.389	-2.43
	(0.008)	(0.011)	(0.02)	(0.224)	(0.13)
Year × Industry × State FE	YES	YES	YES	NO	NO
Firm FE	NO	NO	YES	NO	NO
Number of Technology Groups FE	NO	YES	NO	NO	NO
Year FE	NO	NO	NO	YES	YES
State FE	NO	NO	NO	YES	YES
Industry (NAICS4) FE	NO	NO	NO	YES	YES
Rounded Number of Observations	59,000	59,000	59,000		
Adjusted R^2	0.331	0.332	0.375		
Within R^2	0.242	0.17	0.017		
Rounded Log-Likelihood				-25500	-26000

Notes: Robust standard errors are clustered at the firm level.

Table 12: Knowledge Interdependence and the Number of Startup Formation

DV: IHS-Transformed Number of Startups in Year *t+1*	(1) OLS	(2) OLS	(3) OLS	(4) NB
Log(Knowledge Interdependence)	-0.103	-0.047	-0.02	-0.097
	(0.01)	(0.008)	(0.008)	(0.015)
Log(Number of Technology Groups)	0.174		0.037	0.059
	(0.01)		(0.009)	(0.009)
Log(Firm Size)	0.176	0.175	0.138	0.72
	(0.004)	(0.004)	(0.006)	(0.005)
Diversification	0.285	0.282	0.162	0.213
	(0.026)	(0.025)	(0.03)	(0.028)
Employment Growth	-0.02	-0.019	-0.037	-0.196
	(0.005)	(0.005)	(0.005)	(0.021)
Firm Age	-0.005	-0.004		-0.008
	(0.0005)	(0.0005)		(0.0009)
Constant	-0.524	-0.331	-0.299	-4.506
	(0.02)	(0.022)	(0.03)	(0.133)
Year × Industry × State FE	YES	YES	YES	NO
Firm FE	NO	NO	YES	NO
Number of Technology Groups FE	NO	YES	NO	NO
Year FE	NO	NO	NO	YES
State FE	NO	NO	NO	YES
Industry (NAICS4) FE	NO	NO	NO	YES
Rounded Number of Observations	59,000	59,000	59,000	59,000
Adjusted R^2	0.5	0.511	0.648	
Within R^2	0.338	0.219	0.029	
Rounded Log-Likelihood				-52000

Notes: Robust standard errors are clustered at the firm level.

Table 13: Heterogeneous Effects of Knowledge Interdependence on Potential Founders

DV: Indicator for Founders in Year *t+1*	(1)	(2)	(3)
Log(KI) × Earning Residuals	-0.144		
	(0.042)		
Log(KI) × Earning Rankings		-0.003	
		(0.001)	
Log(KI) × Top 5% Earners			-0.39
			(0.106)
Earning Residuals	0.67		
	(0.056)		
Earning Rankings		0.012	
		(0.001)	
Top 5% Earners			2.555
			(0.138)
Age	0.036	-0.001	0.016
	(0.004)	(0.004)	(0.004)
Age^2	-0.001	<0.001	<0.001
	(0.0004)	(0.0004)	(0.0005)
Job Tenure	-0.061	-0.081	-0.059
	(0.008)	(0.008)	(0.008)
Job Tenure^2	0.002	0.003	0.002
	(0.001)	(0.001)	(0.001)
White	+	+	+
	(***)	(***)	(***)
Male	+	+	+
	(***)	(***)	(***)
Constant	0.382	0.962	0.876
	(0.08)	(0.076)	(0.076)
Rounded Number of Observations	58,860,000	58,860,000	58,860,000
Firm × Year FE	Incld.	Incld.	Incld.
Adjusted R^2	0.0029	0.0028	0.003
Within R^2	0.0002	0.0001	0.0001

Notes: Only signs and significance can be disclosed for variables "White" and "Male" at this stage. Robust standard errors are clustered at the firm level. For exhibition purposes, all numbers in this table have been multiplied by 10^3.

Table 14: Parent Knowledge Interdependence and Startup Survival

DV: Log of the Hazard Ratio	(1)	(2)	(3)	(4)	(5)	(6)
Log(KI)	-0.072	-0.017	-0.073	-0.068	-0.073	-0.06
	(0.013)	(0.023)	(0.013)	(0.013)	(0.013)	(0.014)
Earning Rankings	-0.0005	0.0007				
	(0.0002)	(0.0004)				
Log(KI) × Earning Rankings		-0.001				
		(0.0003)				
Earning Residuals			-0.02	0.037		
			(0.007)	(0.018)		
Log(KI) × Earning Residuals				-0.047		
				(0.014)		
Top 5% Earner					-0.043	0.072
					(0.016)	(0.042)
Log(KI) × Top 5% Earners						-0.097
						(0.033)
Age	-0.021	-0.021	-0.024	-0.024	-0.023	-0.023
	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)
Age^2	0.0003	0.0003	0.0003	0.0003	0.0003	0.0003
	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)
Job Tenure	-0.038	-0.038	-0.038	-0.038	-0.039	-0.039
	(0.007)	(0.007)	(0.007)	(0.007)	(0.007)	(0.007)
Job $Tenure^2$	0.002	0.002	0.002	0.002	0.002	0.002
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Industry Experience	-0.019	-0.019	-0.018	-0.018	-0.019	-0.019
	(0.006)	(0.006)	(0.006)	(0.006)	(0.006)	(0.006)
Log(Firm Employment)	0.001	0.001	0.001	0.001	0.001	0.001
	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)
Log(Number of Technology Groups)	0.034	0.033	0.034	0.034	0.034	0.034
	(0.007)	(0.007)	(0.007)	(0.007)	(0.007)	(0.007)
Firm Age	-0.006	-0.006	-0.006	-0.006	-0.006	-0.006
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Diversification	-0.003	-0.003	-0.002	-0.001	-0.003	-0.003
	(0.022)	(0.022)	(0.022)	(0.022)	(0.022)	(0.022)
Rounded Number of Observations	55,500	55,500	55,500	55,500	55,500	55,500
Gender and Race	Incld.	Incld.	Incld.	Incld.	Incld.	Incld.
Year FE	Incld.	Incld.	Incld.	Incld.	Incld.	Incld.
State FE	Incld.	Incld.	Incld.	Incld.	Incld.	Incld.
Parent Industry FE	Incld.	Incld.	Incld.	Incld.	Incld.	Incld.
Spinout Industry FE	Incld.	Incld.	Incld.	Incld.	Incld.	Incld.
Rounded Log pseudo-likelihood	-354000	-354000	-354000	-354000	-354000	-354000
Rounded No. of Failures	34,000	34,000	34,000	34,000	34,000	34,000

Notes: Regressions are estimated by using the Cox Proportional Hazard Model. The industry of parent firms and spinouts are measured at the 2-digit NAICS level.

Table 15: Knowledge Interdependence and Within Firm Earning Dispersion

DV: Log(Real Earnings)	(1)	(2)	(3)
Log(KI) × Earning Rankings	0.0005		
	(0.0001)		
Log(KI) × Top 5% Earners		0.052	
		(0.01)	
Log(KI) × Bottom 10% Earners			-
			( )
Earning Rankings	0.019		
	(0.0002)		
Top 5% Earners		1.284	
		(0.013)	
Bottom 10% Earners			-
			(***)
Age	0.015	0.088	+
	(0.0004)	(0.001)	(***)
Age^2	-0.0002	-0.0001	-
	(0.000)	(0.000)	(***)
Job Tenure	0.008	0.041	+
	(0.001)	(0.001)	(***)
Job Tenure^2	-0.001	-0.002	-
	(0.000)	(0.000)	(***)
Constant	9.545	8.58	+
	(0.01)	(0.017)	(***)
Rounded # of Observations	58,860,000	58,860,000	-
Gender and Race	Incld.	Incld.	Incld.
Firm × Year FE	Incld.	Incld.	Incld.
Adjusted R^2	0.891	0.627	-
Within R^2	0.818	0.378	-

Notes: In Column (3), *** p<0.001, ** p< 0.01, * p<0.05. Only signs and significance levels can be disclosed for results in Column (3) at this stage. Robust standard errors are clustered at the firm level.

Table 16: Knowledge Interdependence and within-Firm Income Inequality

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Dependent Variable:	Gini	Gini	99-5	99-5	90-10	90-10	75-25	75-25
Log(Knowledge Interdependence)	0.01	-0.001	0.071	0.005	0.034	-0.001	0.016	-0.001
	(0.001)	(0.001)	(0.008)	(0.008)	(0.005)	(0.005)	(0.003)	(0.003)
Log(Number of Technology Groups)	-0.012	0.001	-0.09	0.001	-0.032	-0.004	-0.011	-0.001
	(0.001)	(0.001)	(0.006)	(0.008)	(0.004)	(0.005)	(0.002)	(0.003)
Log(Firm Size)	0.02	0.015	0.161	0.119	0.017	0.013	-0.001	0.004
	(0.001)	(0.001)	(0.002)	(0.005)	(0.001)	(0.003)	(0.001)	(0.002)
Diversification	-0.011	0.011	-0.137	0.089	0.062	0.099	0.065	0.067
	(0.004)	(0.004)	(0.017)	(0.026)	(0.01)	(0.016)	(0.006)	(0.01)
Employment Growth	-0.01	-0.003	-0.078	-0.036	-0.038	-0.02	-0.018	-0.01
	(0.001)	(0.001)	(0.007)	(0.006)	(0.004)	(0.004)	(0.003)	(0.002)
Firm Age	-0.00004		0.0004		-0.002		-0.001	
	(0.0001)		(0.0004)		(0.0002)		(0.0001)	
Constant	0.226	0.238	1.6	1.723	1.323	1.3	0.708	0.674
	(0.003)	(0.005)	(0.013)	(0.024)	(0.008)	(0.015)	(0.005)	(0.009)
Rounded Number of Observations	59,000	59,000	59,000	59,000	59,000	59,000	59,000	59,000
Year × Industry × State FE	Included	Included	Included	Included	Included	Included	Included	Included
Firm FE	No	Included	No	Included	No	Included	No	Included
Adjusted R^2	0.203	0.628	0.2	0.608	0.109	0.579	0.108	0.557
Within R^2	0.096	0.017	0.122	0.019	0.007	0.002	0.004	0.002

Notes: Robust standard errors are clustered at the firm level.

Appendices

Further description of the SDR data set

SDR. The Survey of Doctorate Recipients (SDR) gathers information from a representative sample of individuals who obtained a doctorate in science or engineering from a US PhD-granting institution, conducted every two to three years. It is a longitudinal survey sponsored by the National Center for Science and Engineering Statistics (part of the NSF) that follows respondents until age 76, and at each observation point, new PhD recipients are added while some previously followed individuals are dropped (because of their age or for other reasons). A significant number of individuals were followed over the period from 1993 to 2013, but the sample was almost entirely redrawn in 2015. See http://www.nsf.gov/statistics/srvydoctoratework/ for a detailed description of the target populations and other technical information about SDR.

Startup. The SDR questionnaire classifies individuals as working for startups (new businesses) if they affirmatively answered the question "Did your principal employer come into being as a new business within the past 5 years?" This question was asked in eight surveys, those conducted in 1997, 1999, 2001, 2003, 2010, 2013, 2015, and 2017.

Sample. We use restricted-use SDR data (available through NSF [National Science Foundation] licensing process) from the 1995, 1997, 1999, 2001, 2003, 2006, 2008, 2010, 2013, 2015, and 2017 surveys, all of which are currently available with exception of the 1993 survey, which does not contain some key information (e.g., whether the respondent worked full time). For the analysis of startups, we use eight surveys in which this question was asked, as noted above. We restrict the sample to doctorate recipients who resided in the U.S. and reported being employed full time (defined as 48 weeks or more per year and 30 hours or more per week) with non-zero salaries, whose principal employer was either a private-sector for-profit company or organization or who were self-employed or business owners in their own (incorporated or non-incorporated) businesses. We exclude all retirees and those working part time or not drawing positive salaries, as well as those employed outside the private-for-profit sector. SDR surveys provide weights on individual observations to recover population numbers, and we employ those weights in calculating all the summary statistics and in regression analysis. As mentioned in the main text, using unweighted observations leads to very similar results. Table A1 provides the number of observations in the samples that are used in the main text.

Table A1. Sample number of observations and their distribution across surveys

	Founders		Workers at startups		Workers at established firms	
	Unweighted	Weighted	Unweighted	Weighted	Unweighted	Weighted
1997	177	2,971	966	16,199	7,359	121,885
1999	150	3,111	841	16,493	6,850	137,897
2001	128	2,903	1,192	23,795	6,696	139,754
2003	229	5,433	627	14,770	5,656	131,493
2010	234	6,217	506	13,701	5,964	158,632
2013	208	5,828	455	13,078	5,852	167,431
2015	465	6,746	991	15,802	12,858	193,087
2017	510	6,967	1,289	17,509	14,254	195,057
Total	2,101	40,176	6,867	131,347	65,489	1,245,235

Source: National Science Foundation data.

Work activities: In each survey the respondents were asked to check all work activities that occupied at least 10 percent of their time in a typical week on the job. Table A2 presents the list and more detailed description of all work activities and the number of individuals among founders and workers at established firms who checked the respective boxes. The first four activities listed (basic research, applied research, development, and design) are R&D activities in the definition used in this paper; the next five activities are management activities.

Table A2. Work activities

SDR abbreviation and name	Description of work activity	Founders		Workers at established firms	
		Unweighted	Weighted	Unweighted	Weighted
WABRSH — Basic research	Study aims to gain scientific knowledge for its own sake	419	8,012	16,827	319,956
WAAPRSH— Applied research	Study aims to gain scientific knowledge to meet a recognized need	1,108	21,188	44,185	840,152
WADEV— Development	Using knowledge gained from research for the production of materials and devices	1,043	19,945	39,521	751,469
WADSN— Design	Design of equipment, processes, structures, and models	821	15,699	30,047	571,326
WAACC—Accounting, finance, contracts	Accounting, finance, contracts	1,110	21,226	12,607	239,715
WAEMRL— Human resources	Including recruiting, personnel development, training	686	13,118	17,335	329,615
WAMGMT— Management	Managing or supervising people on projects	1,366	26,121	41,766	794,156
WAQM— Quality management	Quality or productivity management	587	11,225	15,379	292,423
WASALE— Sales	Sales, purchasing, marketing, customer service, public relations	994	19,008	12,931	245,875
WAPROD—Production	Production, operations, maintenance (e.g., chip production, operating lab equipment)	302	5,775	8,729	165,977
WASVC— Professional Services	Professional services (e.g., health care, counseling, financial services, legal services)	952	18,204	10,933	207,885
WACOM— Computer applications	Computer programming, systems, or applications development	736	14,074	25,027	475,874
WATEA— Teaching	Teaching	400	7,649	7,121	135,402
WAOT— Other	Other	141	2,696	4,130	78,530

Source: National Science Foundation data.

Dealing with missing business age data in the 2006 and 2008 surveys

We are missing answers to questions about business age in the 1995, 2006, and 2008 surveys. To obtain a more continuous data series on founders of startups in the 2006 and 2008 surveys we followed Little and Rubin (2019)[38] and estimated the probability of an owner of an incorporated business being a startup founder obtained from those years in which we have information about startups. More specifically, we regressed a dummy equal to one if an owner of an incorporated business was a startup founder and zero otherwise in the 1997, 1999, 2001, 2003, 2010, 2015, and 2017 surveys on a rich vector of individual controls, such as labor experience and its square term, tenure on the job and its square term, demographic characteristics (race, gender, married status, whether there were children in the household, whether the spouse works full time, part time or not at all, and citizenship status), graduate school dummies and education degree class dummies, U.S. state of employment dummies, business/job characteristics (firm size classes, occupation dummies, dummy equal to one if the owner was R&D-focused and zero otherwise, dummy equal to one if the job was related to the education degree and zero otherwise, the number of individuals directly supervised and the number of individuals supervised indirectly) and year dummies. We then assigned an individual in 2006 and 2008 to be a "founder" if the estimated probability of being a founder from the regression above (implemented as a logistic regression although probit and LPM regressions lead to similar outcomes) was 50 percent or more, and to be an owner of an established business otherwise. The correspondence between "founders" predicted from this procedure and actual founders is shown in Table A3.

Table A3.

Actual and predicted founders and established business owners from the logistic regression

		Actual		
		Founder	Established business owner	Total
Predicted	Founder	1,504	407	1,911
	Established business owner	559	6,119	6,678
	Total	2,063	6,526	8,589
Share predicted correctly		0.73	0.94	0.89

Source: Author calculations using weighted National Science Foundation data.

[38] Roderick J.A. Little and Donald B. Rubin, 2019. *Statistical Analysis with Missing Data.* 3rd Edition. John Wiley & Sons, NJ.

As can be seen from Table A3, the predictive power of the logistic regression is quite high as founders are predicted correctly in 73 percent of the cases and established business owners, in 94 percent of the cases. Finally, we compute the implied share of founders in those surveys and use this to fill in the gaps in Figure 1 and in all other Figures in the main text and in the Appendix below. The correspondence between the actual share of founders and the imputed share of founders in all years is presented below, in Figure A1.

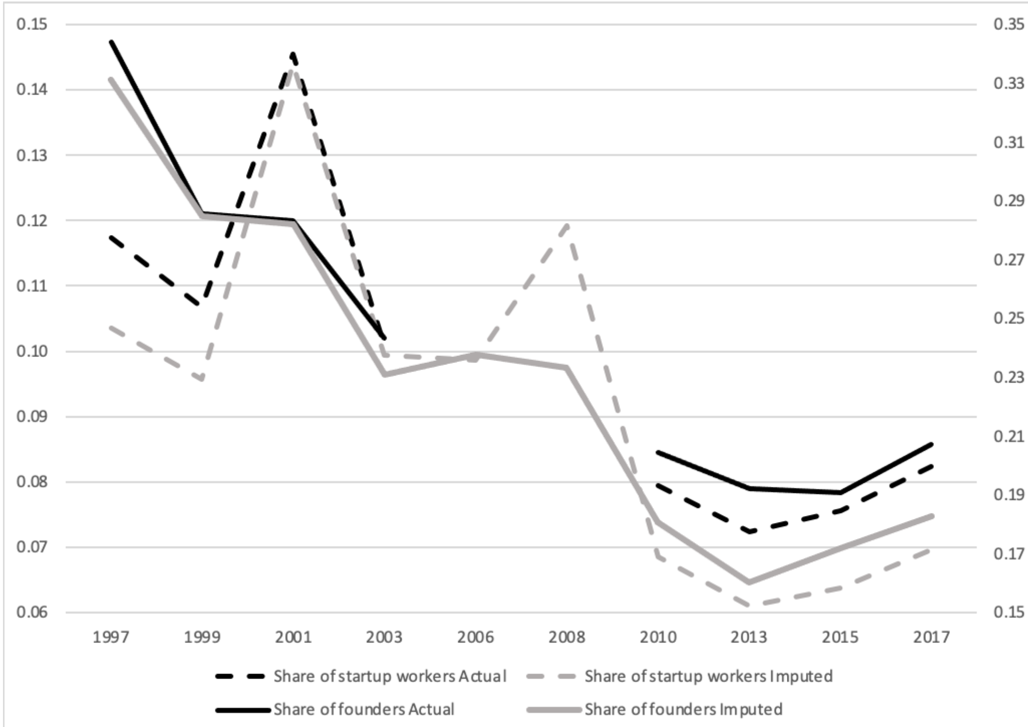
Table A4.

Actual and predicted workers in startups and workers in established firms from the logistic regression

		Actual		
		Workers at startups	Workers at established firms	Total
Predicted	Workers at startups	3,410	2,713	6,123
	Workers at established firms	3,436	61,865	65,301
	Total	6,846	64,578	71,424
Share predicted correctly		0.50	0.96	0.91

We imputed the share of workers in startups using the same procedure as above on the sample of all workers at for-profit firms. Here, to fit the levels, we needed to assign to startups all workers whose estimated probability of working for a startup was 1/3 or higher. As shown in Table A4, we correctly predict who works at an established firm in 96 percent of all cases although we miss about half of the workers who actually work for startups (we can improve this if we assign to startups workers whose estimated probability of working for a startup was even lower than 1/3, but this would lead to overpredicting the levels in Figure A1 below). The overall correspondence between the actual share of workers at startups and the imputed share of such workers in all years is still quite good, as can be seen from Figure A1, although the model somewhat underpredicts these shares in the 2010s. Our predicted data for 2006 and 2008 comport well with an alternative data source. The picture in Figure 3 on p. 7 of Decker et al. (2014) is similar: they find a sharp decline in the number of young firms in the high-tech sector between 2000 and 2012, with some flattening occurring between 2003 and 2008.

Figure A1. Actual and imputed shares of founders and workers at startups

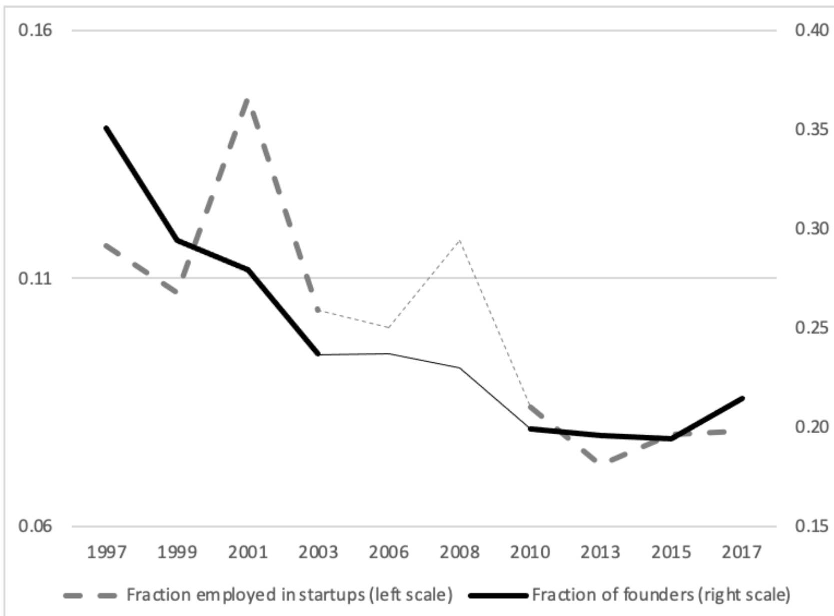


Source: Author calculations using weighted National Science Foundation data.

Universality of the evidence

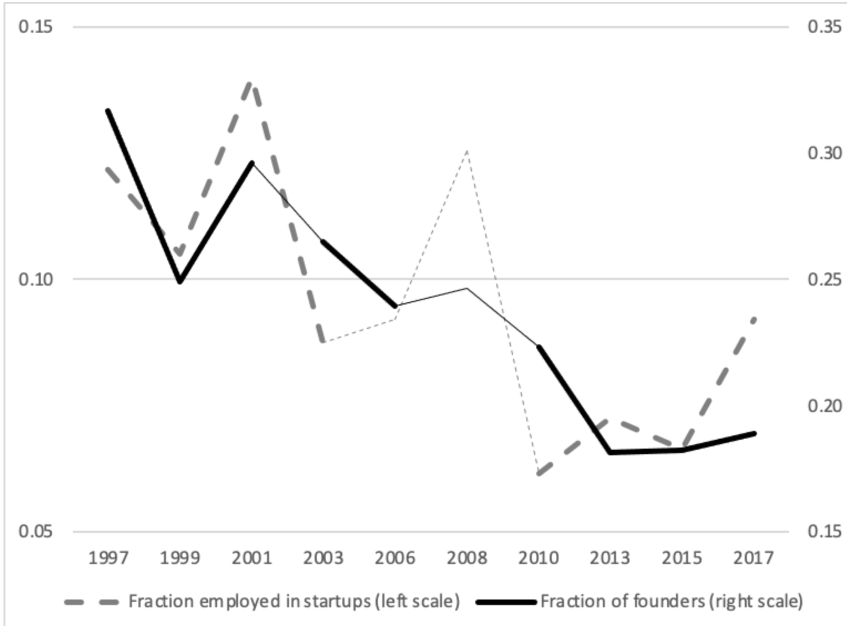
We split the sample by gender and white/non-white race in Figure A2-A5, PhDs employed in California in Figure A6, non-U.S. born PhDs in Figure A7 and PhDs whose work activity included computer-related tasks in Figure A8. The levels are different, but time trends are very similar to those observed in Figure 1. California, in particular, exhibits a decline in the rate of startups as large as that in the rest of the country.

Figure A2. Dynamics of the share of male PhDs employed at or founders of startups



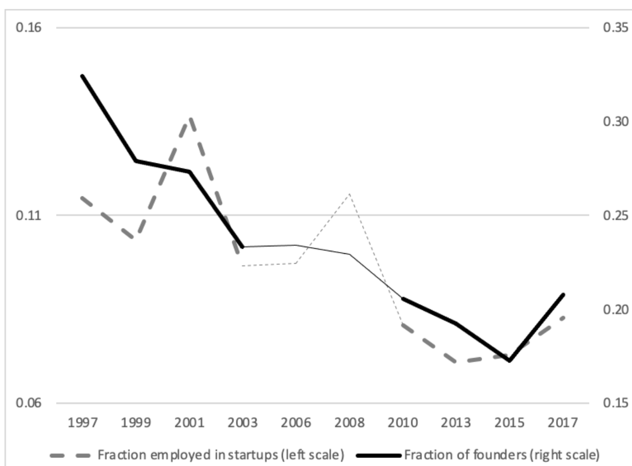
Source: Author calculations using weighted National Science Foundation data. Thinner lines connect imputed data for the 2006 and 2008 surveys. See Appendix.

Figure A3. Dynamics of the share of female PhDs employed at or founders of startups



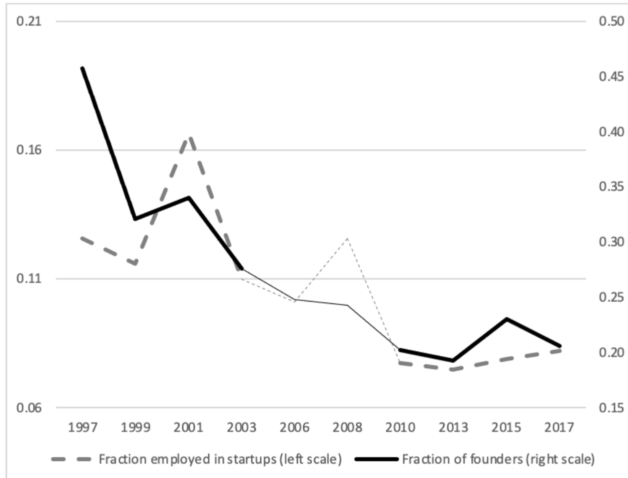
Source: Author calculations using weighted National Science Foundation data. Thinner lines connect imputed data for the 2006 and 2008 surveys. See Appendix.

Figure A4. Dynamics of the share of White PhDs employed at or founders of startups



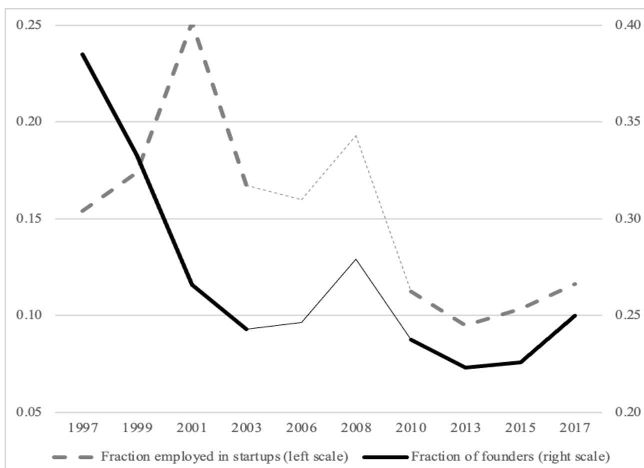
Source: Author calculations using weighted National Science Foundation data. Thinner lines connect imputed data for the 2006 and 2008 surveys. See Appendix.

Figure A5. Dynamics of the share of non-White PhDs employed at or founders of startups



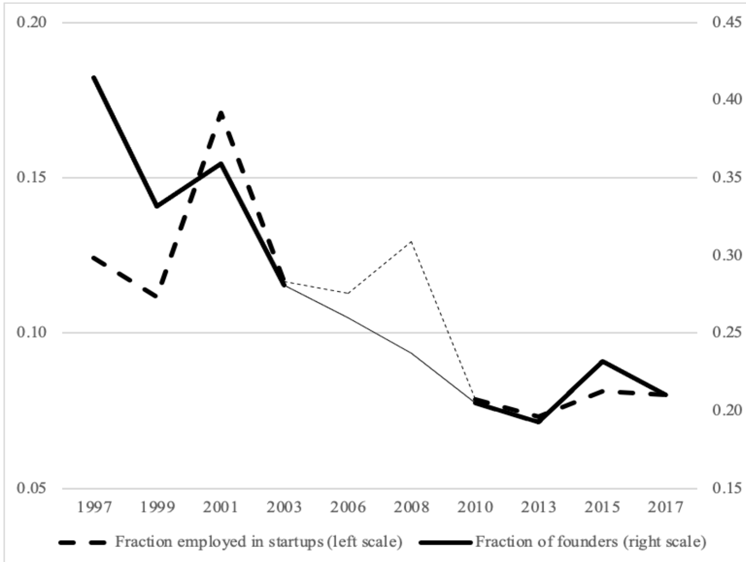
Source: Author calculations using weighted National Science Foundation data. Thinner lines connect imputed data for the 2006 and 2008 surveys. See Appendix.

Figure A6. Dynamics of the share of PhDs residing in California employed at or founders of startups



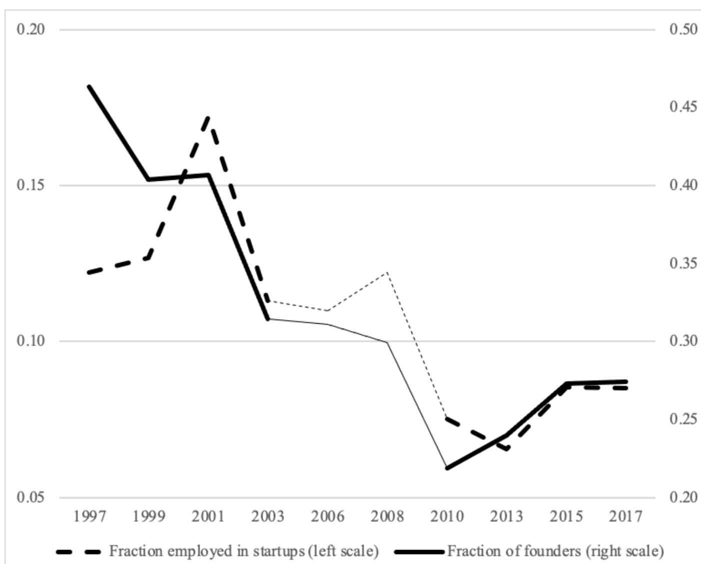
Source: Author calculations using weighted National Science Foundation data. Thinner lines connect imputed data for the 2006 and 2008 surveys. See Appendix.

Figure A7. Dynamics of the share of non-U.S. born PhDs employed at or founders of startups



Source: Author calculations using weighted National Science Foundation data. Thinner lines connect imputed data for the 2006 and 2008 surveys. See Appendix.

Figure A8. Dynamics of the share of PhDs employed at or founders of startups with computer/IT-related tasks

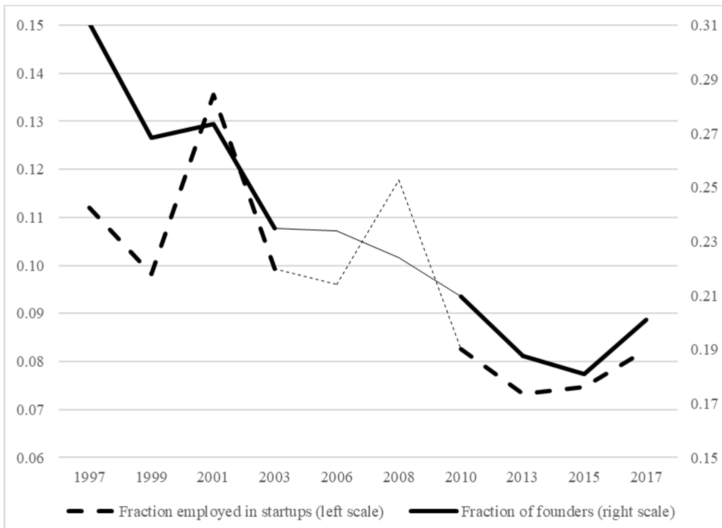


Source: Author calculations using weighted National Science Foundation data. Thinner lines connect imputed data for the 2006 and 2008 surveys. See Appendix.

Computer/IT-related occupations

To make sure that the basic trends toward declining rate of startups (and employment in startups) are not driven by the trends in computer/IT-related occupations, we have reconstructed Figure 1 while excluding all such occupations (computer and information scientists, system analysts, database and network administrators, computer software and hardware engineers, computer programmers, etc.). The results are presented in Figure A9.

Figure A9. Dynamics of the share of PhDs employed in or founders of startups, excluding computer/IT-related occupations



Source: Author calculations using weighted National Science Foundation data. Thinner lines connect imputed data for the 2006 and 2008 surveys. See Appendix.

While the trend in startup founding rates and employment in startups is similar between computer-related and computer-unrelated occupations, the share of doctorate recipients who are employed in computer-related occupations in our data as a whole is steady among incorporated business owners and is actually even slightly increasing over time among those employed in the industry (Table A5). More precisely, there appears to be two periods where the latter fraction is increasing, from 1997-2001 and then once again, from 2008-2017. Significantly, there is no relation between these trends and the trend toward declining startups. That is, there are increasing numbers of Computer Science PhDs graduated, but fewer start new businesses over time.

Table A5. Share of PhDs in computer-related occupations and conducting computer-related tasks.

	Share of PhDs in computer-related occupations in total	
	Industry employment	Incorporated business owners
1995	0.099	0.051
1997	0.128	0.076
1999	0.150	0.070
2001	0.156	0.058
2003	0.138	0.070
2006	0.132	0.068
2008	0.125	0.067
2010	0.140	0.049
2013	0.148	0.062
2015	0.156	0.064
2017	0.167	0.064

Source: Author calculations using weighted National Science Foundation data.

Founders' characteristics

Experience and earnings

As mentioned in the main text, longer work experience prior to founding a startup can be considered to be akin to prolonging the education cycle, and such experience does go up on average for 1.1 years in the raw data. In Table A6 we present the results of a regression analysis where we control for demographics, U.S. state of employment and occupation dummies. The coefficient on the time trend is 0.128, so that over 20 years, the average experience of a founder is estimated to have increased by about 2.6 years (15 percent), more than double of the increase in the raw data.

Table A6. Dynamics of founders' work experience

DV:	Years after PhD
Time	0.128
	(0.037)
Constant	9.58
	(0.350)
Other controls: gender, ethnicity, U.S. state of employment, occupations	Included
Observations	2,101
R-squared	0.196

Estimation method: OLS using weighted National Science Foundation data. Robust standard errors in parentheses.

Interestingly, this does not translate into a much older age for founders. The average age at founding a startup in the raw data was 48.1 years old in 1997 and 48.5 years old in 2017 (an increase of just 0.4 years). If we replace years after PhD by age as the dependent variable in the regression in Table A6, the coefficient on the time trend is 0.096 (statistically significant at the five percent level), which translates into founders in 2017 being 1.9 years older than founders in 1997, an increase of just about four percent. Thus, longer post-PhD experience before launching a startup does not necessarily translate to older age of the founders, which may be related to age at which individuals obtain their PhD degrees declining in the recent decades (see below, Figure A15).

Table A7. Annual real earnings of founders and of workers at established firms (in US $)

	Founders			Workers at established firms		
Survey Year	All	≤13 years after PhD	>13 years after PhD	All	≤13 years after PhD	>13 years after PhD
1997	71,094	72,616	69,958	79,509	66,375	91,846
1999	81,310	81,853	80,909	80,006	68,261	92,460
2001	76,401	80,781	70,676	83,823	73,496	95,248
2003	87,278	80,691	93,398	84,812	74,579	96,936
2010	86,353	71,464	100,780	88,469	76,078	100,849
2013	84,572	63,165	98,453	87,074	73,614	98,341
2015	77,656	61,734	88,001	90,762	76,252	102,370
2017	80,422	57,517	98,690	88,554	75,052	100,890

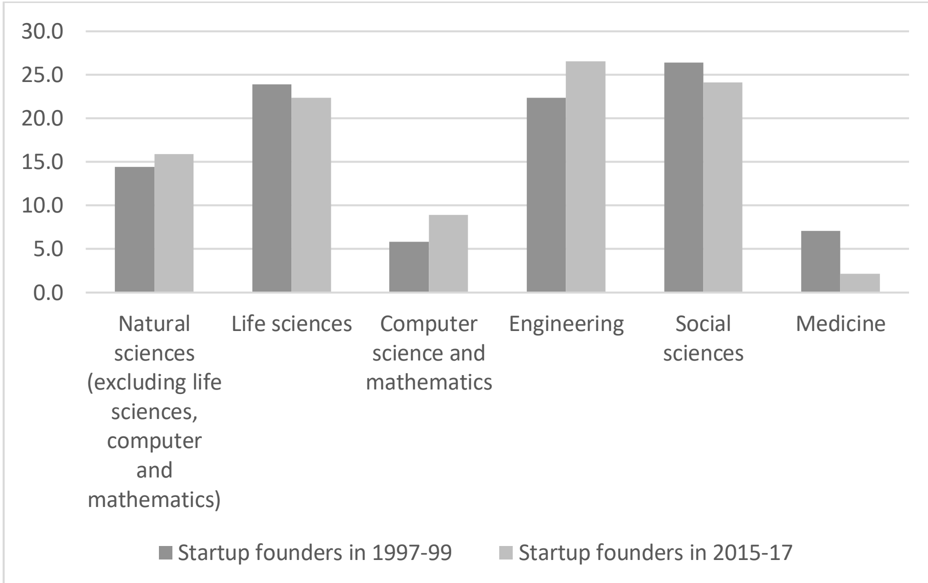
Source: Author calculations using weighted National Science Foundation data. Salary data were adjusted for inflation using consumer price index. Thirteen years after PhD corresponds the median post-PhD experience in the sample.

Table A7 presents the dynamics of inflation-adjusted real earnings of founders and compares these to workers at established firms. Consistent with Figure 2 in the main text and the regression estimation results in Table 3, we can see that earnings of less experienced founders are declining not just in relative but also in absolute terms. By contrast, both more and less experienced workers at established firms increase their earnings over time, as do more experienced founders.

PhD fields

Is there any significant shift of founders' PhD fields over time? Figure A12 compares the fraction of founders with different PhD background in 1997-1999 and in 2015-2017. There is some increase in the fraction of founders with majors in engineering, natural sciences, computer science and mathematics, while the fraction of founders with background in life sciences, social sciences and medicine has declined. But the changes are rather minor and their magnitude cannot account for the sharp decline in the total fraction of startups.

Figure A10. PhD fields of startup founders



Source: Author calculations using weighted National Science Foundation data.

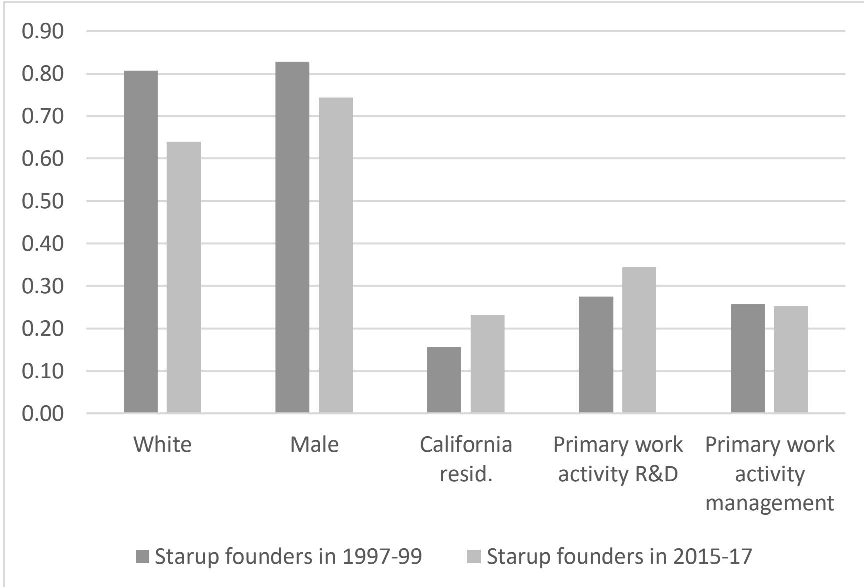
We also wondered if academia may have attracted relatively more PhDs over time but did not find evidence in support of this conjecture. The fraction of PhDs working in academia as opposed to in the for-profit private sector is actually very stable over time. Among all full-time employed individuals residing in the U.S. with non-zero earnings, around 32% of PhDs work in academia and 46% work in the for-profit private sector.

Demographics

We also explored changes in founders' demographics including race, gender, primary work activities and state of residence. Figure A13 shows the fraction of startup founders who are white, male, reside in California, and whether their primary work activity is R&D and management among all founders, respectively. In general, there is smaller share of whites and males among all founders in with a bit more reside in California later years compared to earlier years. Once again, the magnitude of the

changes are not nearly large enough to have a bearing on the overall decline in startups, however. This can also be seen from Figures A2-A7 above which show universal decline in startup rates across various demographics.

Figure A11. Founder demographics



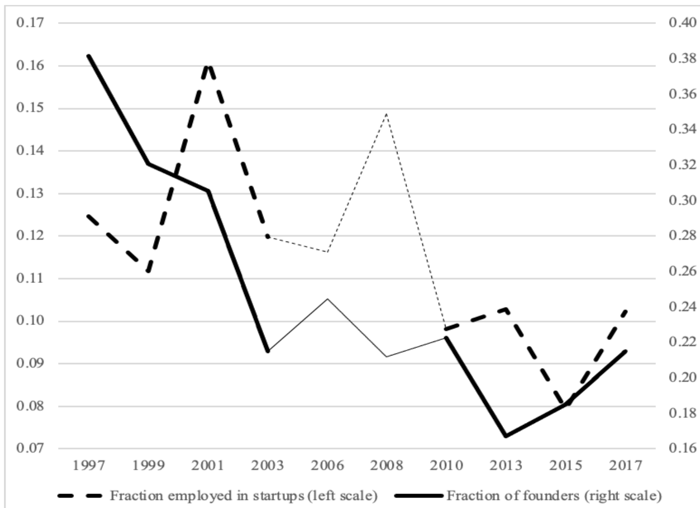
Source: Author calculations using weighted National Science Foundation data.

Life science PhD fields

In our main analysis we control for 72 dummies for various occupation, but many founders list their occupation as "managers," which precludes further analysis. To obtain a proxy for those whose startups are likely to be especially dependent on cumulative knowledge, we use the SDR data on the field of the PhD. Specifically, we define PhDs in life sciences to be those in the 1995-2006 SDR surveys whose fields of major for most recent degree (the MRMED variable until 2001 and the NMRMED variable in 2003 and afterwards) were coded as 216050: Animal sciences, 216060: Food sciences and technology, 216070: Plant sciences, 216080 :OTHER agricultural sciences, 226310: Biochemistry and biophysics, 226320: Biology, general, 226330: Botany, 226340: Cell and molecular biology, 226350: Ecology, 226360: Genetics, animal and plant, 226370: Microbiological sciences and immunology, 226380: Nutritional sciences, 226390: Pharmacology, human and animal, 226400: Physiology and pathology, human and animal, 226410: Zoology, general, 226420: OTHER biological sciences, 236800: Environmental science or studies, 236810: Forestry sciences, and 577240: Bioengineering and biomedical engineering. The share of PhDs in life sciences among all PhDs is stable over the years at about 27-28 percent, while the share of founders with PhDs in life sciences among all founders fluctuates at around 20 percent without a pronounced time trend. We then repeat major parts of our analysis for these PhDs.

In Figure A12 we present the dynamics of the share of founders and workers employed at startups similar to Figure 1 for those with PhDs in life sciences. The picture is similar to Figure 1, but the decline in the share of startups is even more precipitous (a decline of about 50 percent from 1997 to 2017). The decline in the share employed at startups is approximately the same as in Figure 1, suggesting that startup team sizes might be increasing, especially in this category.

Figure A12. Dynamics of the share of life-science PhDs employed at or founders of startups



Source: Author calculations using weighted National Science Foundation data. Thinner lines connect imputed data for the 2006 and 2008 surveys. See Appendix.

Job motivation and satisfaction

As mentioned in the main text, some SDR surveys asked questions about the importance of various factors "when thinking about a job." The listed factors are: (i) salary, (ii) benefits, (iii) job security, (iv) job location, (v) opportunities for advancement, (vi) intellectual challenge, (vii) level of responsibility, (viii) degree of independence, and (ix) contribution to society. The responds are asked to check one of the following boxes for each question: "very important," "somewhat important," "somewhat unimportant," and "not important at all." The questions were asked in the 2001, 2003, 2010, 2013, 2015, and 2017 surveys.

Table A8.
Panel A. Fraction for whom salary is very important

	All startup founders	Young startup founders	All workers in established firms	Young workers in established firms
2001	0.427	0.555	0.526	0.551
2003	0.498	0.544	0.497	0.530
2010	0.562	0.567	0.590	0.625
2013	0.469	0.450	0.595	0.634
2015	0.563	0.541	0.609	0.655
2017	0.528	0.445	0.623	0.665

Source: Author calculations using weighted National Science Foundation data.

Panel B. Fraction for whom the degree of independence is very important

	All startup founders	Young startup founders	All workers in established firms	Young workers in established firms
2001	0.867	0.749	0.620	0.589
2003	0.796	0.762	0.597	0.577
2010	0.820	0.778	0.602	0.585
2013	0.801	0.739	0.588	0.560
2015	0.759	0.755	0.607	0.572
2017	0.788	0.804	0.606	0.571

Source: Author calculations using weighted National Science Foundation data.

The same six surveys also asked the question "Thinking about your principal job, please rate your satisfaction with that job's" various aspects, which, once again, consist of (i) salary, (ii) benefits, (iii) job security, (iv) job location, (v) opportunities for advancement, (vi) intellectual challenge, (vii) level of responsibility, (viii) degree of independence, and (ix) contribution to society. The responds are asked to one answer for each question: "very satisfied," "somewhat satisfied," "somewhat dissatisfied," and "very dissatisfied." Note that this part asks the respondents to think about their current principal job, not just about "a job." Table A9 presents the fraction of those who responded that they were very satisfied with the salary (Panel A) and with the degree of independence (Panel B) for all startup founders and young startup founders (below the median years after PhD) and similarly for PhD workers in established firms.

Table A9.
Panel A. Fraction very satisfied with salary

	All startup founders	Young startup founders	All workers in established firms	Young workers in established firms
2001	0.444	0.526	0.401	0.381
2003	0.435	0.375	0.397	0.377
2010	0.336	0.324	0.378	0.350
2013	0.355	0.330	0.393	0.346
2015	0.413	0.330	0.424	0.379
2017	0.366	0.310	0.425	0.384

Source: Author calculations using weighted National Science Foundation data.

Panel B. Fraction very satisfied with the degree of independence

	All startup founders	Young startup founders	All workers in established firms	Young workers in established firms
2001	0.906	0.896	0.537	0.511
2003	0.882	0.857	0.512	0.477
2010	0.849	0.873	0.489	0.477
2013	0.841	0.787	0.503	0.501
2015	0.868	0.854	0.539	0.513
2017	0.846	0.867	0.545	0.514

Source: Author calculations using weighted National Science Foundation data.

As mentioned in the main text, the decline in the fraction of those satisfied with their salary is most striking among young startup founders, closely following the actual decline in their relative and even absolute earnings we saw in the main text. Their satisfaction with the degree of independence, however, is relatively stable at a very high level, and even after a steep decline in the 2010s, the degree of satisfaction with their salaries is not that much below young paid workers in established firms suggesting that overall, those who choose to launch startups are reasonably happy with their choices.

Type of entrepreneurial opportunity

In the main text, we limited our sample of founders to owners of new incorporated businesses, motivated by the desire to focus on the most innovative and entrepreneurial portion of the sample of business owners and the self-employed. Past research (Braguinsky, Klepper, and Ohyama, 2012) has employed the relation between the job and the highest degree as an alternative proxy for entrepreneurs in these data. SDR surveys contain answers to the question about this relationship, with possible answers being “closely related,” “somewhat related,” and "not related at all." Following established methodology, we constructed a dummy that equals one if the answer was "closely related" and zero otherwise, and we call such business owners/self-employed in new businesses "education-job related" (EJR) founders, regardless of whether their businesses were incorporated. About half the self-employed and business owners and also founders in the sample fall into this category.

Table A10 The number of R&D tasks, the span of control, and the depth of the hierarchy for EJR founders

DV:	IHS-transformed number of R&D tasks	IHS-transformed span of control	IHS-transformed depth of hierarchy
Time	0.0103 (0.002)	0.0061 (0.003)	-0.0015 (0.003)
Experience (years after PhD)	-0.0050 (0.002)	-0.0046 (0.002)	0.0012 (0.003)
Constant	1.108 (0.216)	0.794 (0.446)	0.639 (0.393)
Other controls: gender, ethnicity, occupation, U.S. state of employment	Included	Included	Included
Observations	2,093	2,093	2,093
R-squared	0.478	0.257	0.317

Estimation method: OLS using weighted National Science Foundation data. Robust standard errors in parentheses.

In Table A10 we repeated the estimations of equation (1) in the main text on the sample of EJR founders. The number of R&D tasks exhibits a positive time trend very similar in magnitude to the sample of incorporated startup founders in Table 1. EJR founders have to cope with a number of R&D tasks that are, once again, increasing over time at twice the rate experienced by workers at large, established firms. However, we see that in this case, the time trend is offset by experience (years after PhD), meaning that more experienced EJR founders tend to perform fewer R&D tasks over time. As the founders in the main sample, EJR founders on average do not have recourse to deeper hierarchies, and while the coefficient on time trend in the regression with the span of control as the outcome variable is estimated to be positive, it is only marginally statistically significant.

Table A11. Earnings of EJR founders

DV:	Log (real earnings)
Time	-0.012 (0.004)
Experience (years after PhD)	0.0169 (0.007)
Experience squared	-0.0005 (0.000)
Experience x time	0.0008 (0.000)
IHS-transformed number of R&D tasks	-0.050 (0.035)
Span of control (IHS-transformed # of individuals directly supervised)	0.083 (0.024)
Depth of hierarchy (IHS-transformed # of individuals indirectly supervised)	0.065 (0.021)
Constant	11.064 (0.255)
Other controls: gender, ethnicity, occupation, U.S. state of employment	Included
Observations	2,093
R-squared	0.236

Estimation method: OLS using weighted National Science Foundation data. Robust standard errors in parentheses.

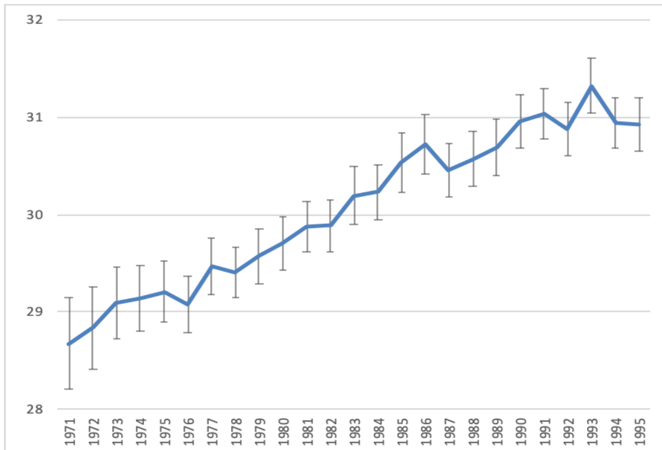
In Table A11 we report the results of the earnings regression on EJR founders. The estimations look similar to those presented for incorporated founders in Table 3 in the main text: a declining time trend offset by the interaction between the time trend and work experience at the mean number of years after PhD; no increase in earnings with the number of R&D tasks (the point estimate is even negative in this sample) and a positive association between earnings and the span of control and the depth of the hierarchy, as in the sample of incorporated founders. Overall, results presented in the main text can be seen to be robust to use of this alternative definition of Schumpeterian firms.

Age at which the PhD is received

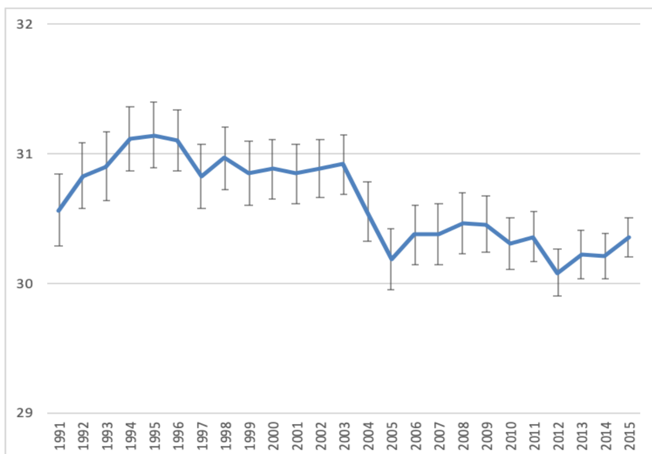
In Figure A15 we plot the dynamics of the age at which individuals received their PhDs. Sample attrition in this case causes those with PhDs received in earlier years to drop out of the sample at a higher rate than those who received their PhDs later, so, rather than pooling all observations, we look at pictures at two points in time, using the 1997 and 2017 surveys. We also limit the sample in each case to those who received their PhD within the 25 years preceding the cutoff year of observations and between ages 20 and 40, to avoid going past the common retirement age.

Figure A15. Average age when the PhD is received

Panel A. 1971-1995 PhD recipients in the 1997 survey



Panel B. 1991-2015 PhD recipients in the 2017 survey



Source: Author calculations using National Science Foundation data. Bars represent 95 percent confidence intervals.

Both Panels A and B (as well as in other years, not shown) show that the average age at which PhDs are received had been increasing until around 1993. Average PhD recipients were about 2.6 years older in 1993 than they were in 1971. Starting in the mid-1990s, however, the average age of receiving a PhD is flat or even declining.

Bibliography

Abowd, J. M., Creecy, R. H., & Kramarz, F. (2002). *Computing Person and Firm Effects Using Linked Longitudinal Employer-Employee Data* (Longitudinal Employer-Household Dynamics Technical Papers No. 2002-06). Center for Economic Studies, U.S. Census Bureau.

Abowd, J. M., Kramarz, F., & Margolis, D. N. (1999). High Wage Workers and High Wage Firms. *Econometrica*, *67*(2), 251–333.

Abraham, K. G., & Taylor, S. K. (1996). Firms' use of outside contractors: Theory and evidence. *Journal of Labor Economics*, *14*(3), 394–424.

Acemoglu, D., & Autor, D. (2011). Chapter 12 - Skills, Tasks and Technologies: Implications for Employment and Earnings. In D. Card & O. Ashenfelter (Eds.), *Handbook of Labor Economics* (Vol. 4, pp. 1043–1171). Elsevier.

Agarwal, R., Braguinsky, S., & Ohyama, A. (2020). Centers of gravity: The effect of stable shared leadership in top management teams on firm growth and industry evolution. *Strategic Management Journal*, *41*(3), 467–498.

Agarwal, R., Campbell, B. A., Franco, A. M., & Ganco, M. (2016). What Do I Take With Me? The Mediating Effect of Spin-out Team Size and Tenure on the Founder–Firm Performance Relationship. *Academy of Management Journal*, *59*(3), 1060–1087.

Agarwal, R., Echambadi, R., Franco, A. M., & Sarkar, M. B. (2004). Knowledge transfer through inheritance: Spin-out generation, development, and survival. *Academy of Management Journal*, *47*(4), 501–522.

Akcigit, U., & Ates, S. T. (2019). *What happened to US business dynamism?* [Working Paper]. National Bureau of Economic Research.

Akcigit, U., & Ates, S. T. (2021). Ten Facts on Declining Business Dynamism and Lessons from Endogenous Growth Theory. *American Economic Journal: Macroeconomics*, *13*(1), 257–298.

Arora, A., Cohen, W. M., Lee, H., & Sebastian, D. (2022). *Invention value, inventive capability and the large firm advantage*. National Bureau of Economic Research.

Åstebro, T., & Tåg, J. (2017). Gross, net, and new job creation by entrepreneurs. *Journal of Business Venturing Insights*, *8*, 64-70.

Autor, D., Dorn, D., Katz, L. F., Patterson, C., & Van Reenen, J. (2020). The Fall of the Labor Share and the Rise of Superstar Firms. *The Quarterly Journal of Economics*, *135*(2), 645–709.

Autor, D. H., Katz, L. F., & Kearney, M. S. (2008). Trends in us wage inequality: Revising the revisionists. *The Review of Economics and Statistics*, *90*(2), 300–323.

Autor, D. H., Levy, F., & Murnane, R. J. (2003). The skill content of recent technological change: An empirical exploration. *The Quarterly Journal of Economics*, *118*(4), 1279–1333.

Ayyagari, M., & Maksimovic, V. (2017). *Fewer and Less Skilled? Human Capital, Competition, and Entrepreneurial Success in Manufacturing* (SSRN Scholarly Article ID 3093443). Social Science Research Network.

Azoulay, P., Jones, B. F., Kim, J. D., & Miranda, J. (2020). Age and high-growth entrepreneurship. *American Economic Review: Insights*, *2*(1), 65–82.
Bagnoli, M., & Bergstrom, T. (2006). Log-concave probability and its applications. In *Rationality and Equilibrium* (pp. 217–241). Springer.
Barth, E., Bryson, A., Davis, J. C., & Freeman, R. (2016). It's where you work: Increases in the dispersion of earnings across establishments and individuals in the United States. *Journal of Labor Economics*, *34*(S2), S67–S97.
Bernard, A. B., & Jensen, J. B. (1995). Exporters, jobs, and wages in us manufacturing: 1976-1987. *Brookings Papers on Economic Activity: Microeconomics*, 1995, 67–119.
Bloom, M. (1999). The performance effects of pay dispersion on individuals and organizations. *Academy of Management Journal*, *42*(1), 25–40.
Bloom, M., & Michel, J. G. (2002). The relationships among organizational context, pay dispersion, and among managerial turnover. *Academy of Management Journal*, *45*(1), 33–42.
Bloom, N., Jones, C. I., Van Reenen, J., & Webb, M. (2020). Are ideas getting harder to find? *American Economic Review*, *110*(4), 1104–1144.
Braguinsky, S., Klepper, S., & Ohyama, A. (2012). High-tech entrepreneurship. *The Journal of Law and Economics*, *55*(4), 869–900.
Braguinsky, S., Mityakov, S., & Liscovich, A. (2014). Direct estimation of hidden earnings: Evidence from Russian administrative data. *The Journal of Law and Economics*, *57*(2), 281–319.
Caliendo, L., Mion, G., Opromolla, L. D., & Rossi-Hansberg, E. (2015). *Productivity and Organization in Portuguese Firms* (No. w21811; p. w21811). National Bureau of Economic Research.
Caliendo, L., & Rossi-Hansberg, E. (2012). The Impact of Trade on Organization and Productivity. *The Quarterly Journal of Economics*, *127*(3), 1393–1467. JSTOR.
Campbell, B. A., Ganco, M., Franco, A. M., & Agarwal, R. (2012). Who leaves, where to, and why worry? Employee mobility, entrepreneurship and effects on source firm performance. *Strategic Management Journal*, *33*(1), 65–87.
Campbell, B. A., Kryscynski, D., & Olson, D. M. (2017). Bridging strategic human capital and employee entrepreneurship research: A labor market frictions approach. *Strategic Entrepreneurship Journal*, *11*(3), 344–356.
Card, D., Heining, J., & Kline, P. (2013). Workplace heterogeneity and the rise of West German wage inequality. *The Quarterly Journal of Economics*, *128*(3), 967–1015.
Carnahan, S. (2017). Blocked but not tackled: Who founds new firms when rivals dissolve? *Strategic Management Journal*, *38*(11), 2189–1430.
Carnahan, S., Agarwal, R., & Campbell, B. A. (2012). Heterogeneity in turnover: The effect of relative compensation dispersion of firms on the mobility and entrepreneurship of extreme performers. *Strategic Management Journal*, *33*(12), 1411–1430.
Castanias, R. P., & Helfat, C. E. (1991). Managerial Resources and Rents. *Journal of Management*, *17*(1), 155–171.

Castanias, R. P., & Helfat, C. E. (2001). The managerial rents model: Theory and empirical analysis. *Journal of Management*, *27*(6), 661–678.

Choi, J., Goldschlag, N., Haltiwanger, J. C., & Kim, J. D. (2021). *Founding teams and startup performance* [Working Paper]. National Bureau of Economic Research.

Chow, M. C., Fort, T. C., Goetz, C., Goldschlag, N., Lawrence, J., Perlman, E. R., Stinson, M., & White, T. K. (2021). *Redesigning the Longitudinal Business Database* [Working Paper]. National Bureau of Economic Research.

Coff, R. W. (1997). Human Assets and Management Dilemmas: Coping with Hazards on the Road to Resource-Based Theory. *The Academy of Management Review*, *22*(2), 374–402.

Davis, S. J., Haltiwanger, J. C., Schuh, S., & others. (1998). Job creation and destruction. *MIT Press Books*, *1*.

De Loecker, J., Eeckhout, J., & Unger, G. (2020). The Rise of Market Power and the Macroeconomic Implications. *The Quarterly Journal of Economics*, *135*(2), 561–644.

Decker, R., Haltiwanger, J., Jarmin, R., & Miranda, J. (2014). The role of entrepreneurship in US job creation and economic dynamism. *Journal of Economic Perspectives*, *28*(3), 3–24.

Decker, R, Haltiwanger, J, Jarmin, R, & Miranda, J. (2016). Declining Business Dynamism: What We Know and the Way Forward. *American Economic Review* 106 (5): 203–207.

Dent, R. C., Pugsley, B. W., Wheeler, H., & others. (2018). *Longitudinal Linking of Enterprises in the LBD and SSL* [Working Paper]. Center for Economic Studies, US Census Bureau.

Densen, P. (2011). Challenges and Opportunities Facing Medical Education. *Transactions of the American Clinical and Climatological Association*, *122*, 48–58.

Dickens, W., & Katz, L. F. (1986). Interindustry wage differences and industry characteristics.

Doctorate Recipients from U.S. Universities 2018 | NSF – National Science Foundation. (n.d.). Retrieved July 15, 2020, from https://ncses.nsf.gov/pubs/nsf20301/report/u-s-doctorate-awards.

Donaldson, D., & Weymark, J. A. (1980). A single-parameter generalization of the Gini indices of inequality. *Journal of Economic Theory*, *22*(1), 67–86.

Dreisigmeyer, D., Goldschlag, N., Krylova, M., Ouyang, W., Perlman, E., & others. (2018). *Building a Better Bridge: Improving Patent Assignee-Firm Links* [Working Paper]. Center for Economic Studies, US Census Bureau.

Elfenbein, D. W., Hamilton, B. H., & Zenger, T. R. (2010). The Small Firm Effect and the Entrepreneurial Spawning of Scientists and Engineers. *Management Science*, *56*(4), 659–681.

Ethiraj, S. K., & Posen, H. E. (2014). Do product architectures affect innovation productivity in complex product ecosystems? In *Collaboration and competition in business ecosystems*. Emerald Group Publishing Limited.

Evans D. & Jovanovic, B. (1989) An Estimated Model of Entrepreneurial Choice

Under Liquidity Constraints. *Journal of Political Economy, 97*(4), 808-827.
Fleming, L., & Sorenson, O. (2001). Technology as a complex adaptive system: Evidence from patent data. *Research Policy*, *30*(7), 1019–1039.
Fleming, L., & Sorenson, O. (2004). Science as a map in technological search. *Strategic Management Journal*, *25*(8–9), 909–928.
Franco, A. M., & Filson, D. (2006). Spin-outs: Knowledge diffusion through employee mobility. *The Rand Journal of Economics*, *37*(4), 841–860.
Freeman, J. (1986). Entrepreneurs as organizational products: Semiconductor firms and venture capital firms. *Advances in the Study of Entrepreneurship, Innovation, and Economic Growth*, *1*(33–52).
Gambardella, A., Ganco, M., & Honoré, F. (2015). Using what you know: Patented knowledge in incumbent firms and employee entrepreneurship. *Organization Science*, *26*(2), 456–474.
Ganco, M. (2013). Cutting the Gordian knot: The effect of knowledge complexity on employee mobility and entrepreneurship. *Strategic Management Journal*, *34*(6), 666–686.
Garicano, L. (2000). Hierarchies and the Organization of Knowledge in Production. *Journal of Political Economy*, *108*(5), 874–904.
Garicano, L., & Rossi-Hansberg, E. (2004). Inequality and the Organization of Knowledge. *American Economic Review*, *94*(2), 197–202.
Garicano, L., & Rossi-Hansberg, E. (2015). Knowledge-based hierarchies: Using organizations to understand the economy. *Economics*, *7*(1), 1–30.
Goldschlag, N. & Akcigit, U. (2022) Employment Dynamics of Patent Inventors, working article, Center for Economic Studies.
Goldschlag, N, & Miranda, J. (2020) Business dynamics statistics of High Tech industries. *J Econ Manage Strat*. 29: 3– 30.
Goldschmidt, D., & Schmieder, J. F. (2017). The rise of domestic outsourcing and the evolution of the German wage structure. *The Quarterly Journal of Economics*, *132*(3), 1165–1217.
Golman, R., & Klepper, S. (2016). Spinoffs and clustering. *The RAND Journal of Economics*, *47*(2), 341–365.
Gompers, P., Lerner, J., & Scharfstein, D. (2005). Entrepreneurial spawning: Public corporations and the genesis of new ventures, 1986 to 1999. *The Journal of Finance*, *60*(2), 577–614.
Goux, D., & Maurin, E. (1999). Persistence of Interindustry Wage Differentials: A Reexamination Using Matched Worker-Firm Panel Data. *Journal of Labor Economics*, *17*(3), 492–533
Grant, R. M. (1996). Toward a knowledge-based theory of the firm. *Strategic Management Journal*, *17*(S2), 109–122.
Guadalupe, M., Li, H., & Wulf, J. (2014). Who Lives in the C-Suite? Organizational Structure and the Division of Labor in Top Management. *Management Science*, *60*(4), 824–844.
Haltiwanger, J., Hathaway, I., & Miranda, J. (2014). Declining business dynamism in the US high-technology sector. *Available at SSRN 2397310*.
Haltiwanger, J., Jarmin, R. S., Kulick, R., & Miranda, J. (2016). High growth young

firms: Contribution to job, output, and productivity growth. In *Measuring entrepreneurial businesses: Current knowledge and challenges* (pp. 11–62). University of Chicago Press.
Hamilton, B. H. (2000). Does Entrepreneurship Pay? An Empirical Analysis of the Returns to Self-Employment. *Journal of Political Economy*, *108*(3), 604–631.
Heckman, J. J., & Singer, B. (2017). Abducting economics. *American Economic Review*, *107*(5), 298–302.
Hellmann, T. (2007). When do employees become entrepreneurs? *Management Science*, *53*(6), 919–933.
Hombert, J., & Matray, A. (2019). *Technology Boom, Labor Reallocation, and Human Capital Depreciation* (SSRN Scholarly Article ID 3496625). Social Science Research Network.
Hopenhayn, H., Neira, J., & Singhania, R. (2018). *From Population Growth to Firm Demographics: Implications for Concentration, Entrepreneurship and the Labor Share* (Working Article No. 25382). National Bureau of Economic Research.
Howell, S. T., & Brown, J. D. (2020). *Do Cash Windfalls Affect Wages? Evidence from R&D Grants to Small Firms* (SSRN Scholarly Article ID 3530688). Social Science Research Network.
Hurst, E., & Pugsley, B. W. (2011). What Do Small Businesses Do. *Brookings Articles on Economic Activity*, 73–118. JSTOR.
Jones, B. F. (2009). The burden of knowledge and the “death of the renaissance man”: Is innovation getting harder? *The Review of Economic Studies*, *76*(1), 283–317.
Karahan, F., Pugsley, B., & Şahin, A. (2019). *Demographic Origins of the Startup Deficit* (Working Article No. 25874). National Bureau of Economic Research.
Katz, L. F., & Krueger, A. B. (2019). The rise and nature of alternative work arrangements in the United States, 1995–2015. *ILR Review*, *72*(2), 382–416.
Katz, L. F., & Summers, L. H. (1989). Industry rents: Evidence and implications. *Brookings Papers on Economic Activity. Microeconomics*, 1989, 209–290.
Klepper, S. (2016). *Experimental Capitalism: The Nanoeconomics of American High-Tech Industries*. Princeton University Press.
Klepper, S., & Thompson, P. (2010). Disagreements and intra-industry spinoffs. *International Journal of Industrial Organization*, *28*(5), 526–538.
Lazear, E. P. (2005). Entrepreneurship. *Journal of Labor Economics*, *23*(4), 649–680.
Levine, R., & Rubinstein, Y. (2017). Smart and illicit: who becomes an entrepreneur and do they earn more?. *The Quarterly Journal of Economics*, *132*(2), 963-1018.
Lerner, J., & Nanda, R. (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.
Little, R. J. A., & Rubin, D. B. (2019). *Statistical Analysis with Missing Data*. John Wiley & Sons.
Lucas, R. E. (1978). On the size distribution of business firms. *The Bell Journal of Economics*, 508–523.

Luo, J., & Wood, K. L. (2017). The growing complexity in invention process. *Research in Engineering Design*, *28*(4), 421–435.

National Science Foundation, National Center for Science and Engineering Statistics. 2019. *Doctorate Recipients from U.S. Universities: 2018. Special Report NSF 20-301*. Alexandria, VA. Available at https://ncses.nsf.gov/pubs/nsf20301/.

National Venture Capital Association. 2019. *2019 Yearbook*. Washington, DC: National Venture Capital Association.

Nickerson, J. A., & Zenger, T. R. (2004). A knowledge-based theory of the firm—The problem-solving perspective. *Organization Science*, *15*(6), 617–632.

Ohyama, A. (2015). Entrepreneurship and Job-relatedness of Human Capital. *Economica*, *82*(328), 740–768.

Ouimet, P., & Zarutskie, R. (2014). Who works for startups? The relation between firm age, employee age, and growth. *Journal of Financial Economics*, *112*(3), 386–407.

Peirce, C. S. (1997). *Pragmatism as a principle and method of right thinking: The 1903 Harvard lectures on pragmatism*. Suny Press.

Pillai, S. D., Goldfarb, B., & Kirsch, D. (2021). *Using historical methods to improve abduction and inference to the best explanation in strategy*. HBS Mimeo.

Rivkin, J. W. (2000). Imitation of complex strategies. *Management Science*, *46*(6), 824–844.

Rivkin, J. W., & Siggelkow, N. (2002). Organizational sticking points on NK landscapes. *Complexity*, *7*(5), 31–43.

Roach, M., & Sauermann, H. (2015). Founder or Joiner? The Role of Preferences and Context in Shaping Different Entrepreneurial Interests. *Management Science*, *61*(9), 2160–2184.

Roach, M., & Skrentny, J. (2019). Why foreign STEM PhDs are unlikely to work for US technology startups. *Proceedings of the National Academy of Sciences*, *116*(34), 16805–16810.

Sakakibara, M., & Balasubramanian, N. (2020). Human capital, parent size, and the destination industry of spinouts. *Strategic Management Journal*, *41*(5), 815–840.

Schumpeter, J. A. (1911). *The theory of economic development*. Cambridge, MA: Harvard University Press.

Schumpeter, J. A. (1942). *Socialism, capitalism and democracy*. Harper and Brothers.

Shani, A. B., Coghlan, D., & Alexander, B. N. (2020). Rediscovering abductive reasoning in organization development and change research. *The Journal of Applied Behavioral Science*, *56*(1), 60–72.

Shaw, J. D., Gupta, N., & Delery, J. E. (2002). Pay dispersion and workforce performance: Moderating effects of incentives and interdependence. *Strategic Management Journal*, *23*(6), 491–512.

Song, J., Price, D. J., Guvenen, F., Bloom, N., & Von Wachter, T. (2019). Firming up inequality. *The Quarterly Journal of Economics*, *134*(1), 1–50.

Sorenson, O., Rivkin, J. W., & Fleming, L. (2010). Complexity, networks and knowledge flows. In *The Handbook of Evolutionary Economic Geography*. Edward Elgar Publishing.

Sorkin, I. (2018). Ranking firms using revealed preference. *The Quarterly Journal of Economics*, *133*(3), 1331–1393.
Starr, E., Ganco, M., & Campbell, B. A. (2018). Strategic human capital management in the context of cross-industry and within-industry mobility frictions. *Strategic Management Journal*, *39*(8), 2226–2254.
Segal, L. M., & Sullivan, D. G. (1997). The growth of temporary services work. *Journal of Economic Perspectives, 11*(2), 117–136.
Stern, S. (2004). Do Scientists Pay to Be Scientists? *Management Science*, *50*(6), 835–853.
Tåg, J., Åstebro, T., & Thompson, P. (2016). Hierarchies and entrepreneurship. *European Economic Review*, *89*, 129–147.
Takii, K. (2003). Prediction ability. *Review of Economic Dynamics*, *6*(1), 80–98.
Tate, G., & Yang, L. (2015). The bright side of corporate diversification: Evidence from internal labor markets. *The Review of Financial Studies*, *28*(8), 2203–2249.
Toh, P. K., & Ahuja, G. (2022). Integration and appropriability: A study of process and product components within a firm's innovation portfolio. *Strategic Management Journal*, *43*(6), 1075–1109.
Tohmé, F., & Crespo, R. (2013). Abduction in economics: A conceptual framework and its model. *Synthese*, *190*(18), 4215–4237.
Wang, H., Zhao, S., & Chen, G. (2017). Firm-specific knowledge assets and employment arrangements: Evidence from CEO compensation design and CEO dismissal. *Strategic Management Journal*, *38*(9), 1875–1894.
Weil, D. (2014). The fissured workplace. *The fissured workplace.* Harvard University Press.
Wuchty, S., Jones, B. F., & Uzzi, B. (2007). The increasing dominance of teams in production of knowledge. *Science*, *316*(5827), 1036–1039.
Yayavaram, S., & Ahuja, G. (2008). Decomposability in knowledge structures and its impact on the usefulness of inventions and knowledge-base malleability. *Administrative Science Quarterly*, *53*(2), 333–362.
Zenger, T. R. (1992). Why Do Employers Only Reward Extreme Performance? Examining the Relationships Among Performance, Pay, and Turnover. *Administrative Science Quarterly*, *37*(2), 198–219.