Entrepreneurial Spillovers from Corporate R&D

Tania Babina and Sabrina T. Howell*

September 19, 2022

Abstract

This paper offers the first study of how changes in corporate R&D investment affect labor mobility. We show that increases in firm R&D spur employee departures to join startups' founding teams. This appears to reflect employees taking ideas, skills, or technologies to startups that are created through the R&D process but are not especially valuable to the R&D-investing firm. The employee-founded startups tend to be outside the R&D-investing employer's industry, suggesting that the underlying ideas would impose diversification costs on the R&D-investing firm. The startups are also more likely to be VC-backed, high-tech, and high-wage, pointing to substantial spillover benefits.

*Tania Babina is an Assistant Professor, Finance Department, Columbia Business School, Columbia University, New York, 10027, USA, tania.babina@gsb.columbia.edu. Sabrina Howell (corresponding author) is an Assistant Professor, Finance Department, Stern School of Business, NYU & NBER, New York, 10012, USA, sabrina.howell@nyu.edu. This paper relies on US Census Bureau data. Access expired in March 2019, preventing us from doing additional analysis unless we have a request for revision from a journal. We thank seminar participants at the GSU CEAR Conference, Texas Finance Festival, Chicago Booth, UBC, SoCal PE Conference, Columbia GSB, the Federal Reserve Board, the IMF, Duke, Drexel, Toulouse School of Economics, US Census Bureau, NY Fed, Yale Junior Conference, NBER Corporate Finance, Stanford Financing of Innovation Summit, LBS Corporate Finance Conference, Stanford SIEPR, and Washington Summer Finance Conference. We are grateful to Belinda Chen for excellent research assistance, and Chris Galvan and Shirley Liu for data assistance. Tania Babina was a Special Sworn Status researcher of the US Census Bureau at the NY Research Data Center. Any opinions and conclusions expressed herein are those of the author(s) and do not represent the views of the Census Bureau. No confidential information is disclosed. This research uses data from the Census Longitudinal Employer Household Dynamics Program, which was partially supported by the following NSF Grants SES-9978093, SES-0339191 and ITR-0427889; National Institute on Aging Grant AG018854; and grants from the Alfred P. Sloan Foundation. The Kauffman Foundation funded Sabrina Howell's work on this project.

1 Introduction

Corporate research and development (R&D) investment generates new knowledge, technology, and skills. Some of these outputs benefit other firms through spillovers, which are central to economic growth (Romer, 1990, Atkeson and Burstein, 2010, Bloom et al. 2013). However, the mechanisms behind spillovers are less well-established. While the patent literature demonstrates knowledge diffusion through citations and inventor mobility (Griliches 1992, Jaffe et al. 1993), other mechanisms are likely also at work. Many innovative firms do not patent. Furthermore, patents capture only the subset of innovation outputs that are contractible and over which the firm has chosen to establish property rights (Sampat 2018, Garcia-Macia et al. 2019). Another important potential channel is through labor mobility. Employees, whose human capital is inalienable and portable (Aghion and Tirole 1994), are a means for new knowledge and skills to leave the firm, especially to startups.

Consider two examples. First, in 1894 Henry Ford left Thomas Edison's Illuminating Company, which constructed electrical generating stations, to launch his own venture. Two years later, he produced the first Ford Quadricycle with the help of a local angel investor and employing mechanical and electrical engineering advances from Edison (Glaeser 2011). Second, in the 1990s, Michael Rosenfelt worked for the computer memory company Micron Electronics, where he helped to revitalize its PC business. He left in 1999 to found Powered Inc., a successful online education company that exploited marketing innovations from Micron.[1] These anecdotes highlight how employees may take their employer firm's innovation outputs and exploit them in a new firm, building products that are far from the employer firm's core focus.

We hypothesize that R&D investment may generate skills or ideas that are specifically valuable for startups (Hart and Moore, 1990; Gromb and Scharfstein, 2002). However, there may be no effects if, for example, firms can commit to long-term wage contracts with their workers and fully internalize productivity spillovers (Heggedal et al. 2017). Alternatively, effects may only occur along other dimensions such as retention, layoffs, or mobility to other firms.

This paper's main contribution is to provide the first direct evidence on the relationship between

[1] For evidence of success, see here, here, and here.

firm R&D investment and worker mobility. To do this, we construct panel data combining worker-firm US Census micro data with US public firm R&D investment. This approach allows us to examine whether changes in firm R&D investment predict labor mobility–and potential spillovers–to other firms, focusing on startups, which are crucial for growth.[2] We use a regression model with firm, state-year, and granular industry-year fixed effects, as well as time-varying firm characteristics and local market-level R&D controls to estimate the effect of R&D within publicly traded firms. This strategy has advantages over the cross-sectional approaches used in the literature. To fix ideas with an example, our empirical strategy compares employee departures within Microsoft between periods of higher and lower R&D, rather than comparing employee departures between Microsoft and Walmart.

We find robust effects of within-firm R&D increases on employee departures to entrepreneurship, an intuitive result because startups are known to be important conduits for commercializing new ideas. Specifically, a 100 percent increase in R&D (about one standard deviation) predicts an 8.4 percent increase in employee entrepreneurship. We use the term "entrepreneurship" in a broad sense to mean the founding team of a new firm; the group most likely to contribute ideas and crucial skills to the startup.[3] Over the course of the sample, above- relative to below-median R&D changes within firms yield 8,291 additional employee-founded startups, which is 7.7 percent of all employee-founded startups in the data. As we expect, the effect is higher among high-tech establishments of employer (i.e., parent) firms (e.g. Amazon's headquarters rather than its warehouses). The effect is not driven by recent hires who might have been hired because of a new R&D project and is robust to alternative measures of both R&D and entrepreneurship, including the number of startups founded by recently departed employees. We find no effects on departures to other incumbent firms, on unemployment, or on employee retention.

We find cross-sectional support for a mechanism in which employees take ideas, skills, or technologies to startups that are created though the R&D process but are not especially valuable to

[2] Entrepreneurs play a crucial role in prominent theoretical explanations for economic growth, including Schumpeter (1911), Lucas (1978), and Baumol (1990). Relative to incumbent firms, new firms have faster productivity and employment growth. This literature includes Kortum and Lerner (2000), Foster et al. (2008), Gennaioli et al. (2012), Haltiwanger et al. (2013) Decker et al. (2014), and Glaeser et al. (2015).

[3] Our main entrepreneurship outcome is the share of an establishment's employees who depart and are among the top five earners of a firm founded within three years. This captures founders and early employees. Similar variables are used in Kerr and Kerr (2017), Babina (2020), and Azoulay et al. (2020).

the parent firm. This mechanism has two premises. First, the innovation process is serendipitous, producing unforeseen outputs. Second, innovation effort is hard to contract on ex-ante, and hard to verify ex-post (Grossman and Hart 1986, Aghion and Tirole 1994). Contending with these frictions, the firm may opt not to pursue all good innovations, enabling employees to take some outside the firm. We expect this mechanism to be particularly salient when the idea is less valuable to the parent firm, which may more often be the case when the idea is far from the firm's core focus and would impose diversification costs. Diversification has been shown to negatively affect productivity and innovation.[4] While transaction costs are lower within the boundary of the firm (Atalay et al. 2019), there is relatively less benefit to locating within the same firms assets that are not complementary (Williamson 1975, Hart and Moore 1990). Consistent with this channel, we find that R&D-induced startups tend to be in a different broad industry from the parent firm. We also show using supply chain relationships that R&D-induced startups are more likely to draw inputs from a broader array of supplier industries.

To access the best talent and induce optimal effort, firms may permit employees to depart with R&D outputs. A permissive approach to employee entrepreneurship is likely most feasible if the lost innovations are not ones the firm is able to sell or develop. Indeed, we find that contractible R&D outputs over which the firm does establish explicit property rights—measured by patents—do not yield employee-founded startups. It is usually impossible for firms to sell non-patentable ideas to other firms (Akcigit et al. 2016). This may contribute to employees' ability to take some ideas outside the firm to startups. Alternative channels may play a role, but they have less support in the data. For example, the evidence is inconsistent with a mechanism in which employees steal ideas that the firm values.

In addition to the diversification channel, we expect larger effects when an idea is especially risky but potentially high growth.[5] These ideas benefit from the high-powered incentives that exist in small, focused firms (Rhodes-Kropf and Robinson 2008, Phillips and Zhdanov 2012) and are also more conducive to spillovers and growth. Consistent with this mechanism, higher parent R&D is strongly associated with venture capital (VC) backing among employee-founded startups. We find that 2% of startups in our sample ever receive VC funding, which is 18 times the national average. Moreover, we

[4] See Mullainathan and Scharfstein (2001), Schoar (2002), Maksimovic and Phillips (2002), and Seru (2014).

[5] This is due to contracting frictions, as described in Gromb and Scharfstein (2002), Robinson (2008), and Frésard et al. (2020).

document that R&D is by far the strongest predictor of whether an employee-founded startup receives VC among the dozens of parent firm characteristics that we observe. Also, R&D-induced startups pay higher wages on average and are more likely to be incorporated, in high-tech sectors, and exit (fail or be acquired). Therefore, the effect appears to be driven by risky, new-to-the-world ideas, rather than Main Street-type businesses. In sum, the types of ideas that employees take to entrepreneurship seem to be those that benefit from focused, high-powered incentives and that are not especially complementary with the firm's existing activities. R&D-induced startups tend to be the types that are an important source of economic spillovers and growth.

Entrepreneurial spillovers from R&D could be costly to the parent firm, and these costs may be priced into the labor contract. Importantly, there would be no spillover effects from R&D-induced employee entrepreneurship only if the startups are wholly owned spinouts and parent firms fully internalize their benefits. We present evidence that this is not the case, because parent firms very rarely invest in or acquire these startups.[6] Therefore, the effect appears to be a "spillover" in the sense of being a benefit of one firm's R&D that accrues to another firm. We do not assess the welfare effects of R&D-induced startups, but our finding suggests greater corporate underinvestment in R&D relative to the social optimum, which would include the social and private benefits of R&D-induced startups.

Our results also shed light on how large firms and entrepreneurs interact. One narrative is that large firms negatively impact entrepreneurship (Glaeser et al. 2015), for example because startups find it difficult to compete for talent against large, well-resourced firms.[7] We present a different perspective, in which large firms contribute to startup formation when some of their innovation output spills over into employee-founded firms. Indeed, the effect increases when the parent firm is larger. Of course, we cannot speak to the general equilibrium effect of a large firm on overall local entrepreneurship.

A range of robustness tests and an IV specification offer support for a causal interpretation of the main results. The main concern with the OLS estimates is that a technological opportunity may jointly engender parent R&D and employee entrepreneurship, leading to upward bias. In this case, we

[6] We cannot rule out parents sometimes having licensing or investment contracts with R&D-induced startups.

[7] One example is the concern among tech entrepreneurs in the Washington, DC area about the arrival of Amazon's HQ2. One entrepreneur said, "Amazon is just going to hire our people...and technical workers who are skilled will become more scarce and more expensive." See here.

expect that controlling for variables known to be correlated with technology investment opportunities, even imperfectly, will attenuate the relationship between R&D and employee departures to entrepreneurship. To test this, we add time-varying firm variables to the within-firm estimates, including patenting activity, sales growth, profitability, Tobin's Q, and cash. Second, we include industry-by-year and state-by-year fixed effects, which should correlate with industry or location technology shocks.[8] Including these controls does not attenuate the main point estimate. Furthermore, a technology shock channel should also increase employee departures to incumbent firms, but we only observe reallocation to startups. Finally, the technology shock channel predicts that R&D-induced startups will be closely related to the R&D-performing parent. Instead, as discussed above they tend to be in a different industry from the parent firm. In sum, while our estimation cannot completely rule out omitted variable bias, the data do not support a channel in which a technological opportunity leads to both higher parent R&D and more employee reallocation to new firms.

To confirm the effect, we instrument for R&D using changes in state and federal R&D tax credits, which alter the firm's user cost of R&D. We closely follow Bloom et al. (2013). The instruments do not permit us to affirmatively establish causality as though R&D were randomly allocated across firms, but they offer a useful robustness test, as in König et al. (2019). The instruments satisfy the relevance condition and are likely to satisfy the exclusion restriction, which is that tax credits can affect employee entrepreneurship only through the employer's R&D.[9] Like the OLS analysis, the instrumental variables (IV) analysis finds that R&D has a strong positive effect on employee-founded startups, but no effect on retention or departures to incumbents.

This paper contributes to the literature on labor mobility and innovation (e.g. Kim and Marschke 2005, Kerr and Lincoln 2010, Balsvik 2011, Kaiser et al. 2015, Serafinelli 2019). By focusing on innovation inputs and labor reallocation across firms, this paper also extends the empirical literature on R&D spillovers, which includes Jones and Williams (1998) and Kerr and Kominers (2015). One line of research focuses on the role of competitive incentives (Atkeson and Burstein 2019, López and Vives 2019), another on spillovers within firms (Herkenhoff et al. 2018, Bilir and Morales 2020),

[8] We use SIC 3-digit industries, which are already quite granular, but find similar effects with 4-digit industries.

[9] For example, see Bankman and Gilson (1999), Curtis and Decker (2018), Lucking (2018), and Lucking et al. (2019).

another on financing frictions (Bai et al. 2018, Babina 2020, Babina et al. 2022a), and a large strand on spatial aspects of spillovers (Jaffe et al. 1993, Griffith et al. 2006, Kantor and Whalley, 2014, Ganguli et al. 2020). To our knowledge, this paper is the first to measure how changes in corporate R&D investment affect labor reallocation to other incumbent firms and to startups as a potential channel of knowledge spillovers. Since our effect is local – with 88% of the employee-founded startups in the same state as the former employer – it offers one micro-foundation for agglomeration advantages based on knowledge spillovers (Audretsch and Feldman 2004, Combes and Duranton 2006). This is relevant not only for understanding the formation of U.S. innovation hubs in our dataset–such as those in North Carolina and Texas–but also clusters in other countries, such as the "Third Italy" and Germany's Baden-Wuerttemberg.

Our finding helps explain why high-growth startup founders are often former employees of large incumbent firms (Gompers et al. 2005, Babina 2020). We offer corporate R&D as a new source for where ideas and human capital for the entrepreneurial sector come from.[10] Contributing to the debate about declining dynamism (Decker et al. 2016, Guzman and Stern 2017), we find that on one margin—through employees—corporate R&D increases high-growth entrepreneurship. This is particularly relevant as corporate R&D has increased in recent decades while other sources, such as federal spending, have declined (Goldfarb et al. 2009, Van Reenen 2020, Babina et al. 2022b).

Other related work on knowledge diffusion has emphasized the importance of inventor networks and collaboration among scientists (Borjas and Doran 2015, Zacchia 2019). Our results offer one mechanism for how such networks can emerge: firm investments in R&D can induce human capital to move from one firm to another. In our case, R&D outputs move via an employee from a large firm to a new startup that is typically in a different industry. Employee entrepreneurship to other sectors is a channel for knowledge spillovers and technological interconnections across sectors.

[10] Work on this topic, including incumbent firm spinoffs, includes Klepper (2001), Agarwal and Shah (2014), Aghion and Jaravel (2015), and Akcigit and Ates (2019).

2 Conceptual Framework

We hypothesize that some of the new growth options emerging from innovation investment at larger firms are reallocated via employees to startups. Contracting and verification frictions imply that growth options (e.g., ideas, knowledge or application of new skills/ technologies) emerging from R&D may sometimes be optimally developed outside of the firm's boundaries in new, standalone firms, because there are benefits to allocating residual rights of control to the party that performs innovation (Aghion and Tirole, 1994). If the employee is responsible for the investment necessary to incubate an idea, he may exert optimal effort only in his own firm. Frésard et al. (2020) model these frictions and conclude that control rights should be allocated to stand-alone firms in especially R&D-intensive industries and when the innovation is as yet unrealized; that is, when innovation requires more unverifiable effort. This does not mean that it is optimal to assign control rights to the innovators within the firm, or that explicit ideas developed with company assets are not the firm's IP. Instead, it means that the firm may not prioritize keeping all the ideas that serendipitously emerge from R&D and the employees associated with them in-house.

The firm may permit employee entrepreneurship to continue hiring the best talent. Suppose that corporate innovation requires partially unobservable employee effort, and sometimes yields some outputs that are more valuable for the employee to pursue in a standalone firm than in the R&D-performing firm. The firm can (a) give nothing to the employee, and perhaps try to sell the idea (although as Akcigit et al. (2016) point out, most ideas are not patentable and so this is often impossible); (b) try to share the idea with the employee, perhaps through a joint venture; or (c) permit the employee to leave the firm with the idea. If the firm can commit to allowing employees to "walk away" with ideas that are relatively less valuable to the firm, this reward may help the firm to hire the right talent (the participation constraint) and induce optimal effort (the incentive compatibility constraint). Along these lines, Kondo et al. (2021) model the importance of trust in the firm-inventor relationship; to maintain its reputation, the firm enables the inventor to pursue some of his own ideas.

We expect that if this mechanism is at play, R&D-induced employee-founded startups will be more likely to be risky and potentially high-growth. Agency frictions are magnified when an idea is riskier, making high-risk, high-reward growth options more often best located outside the firm boundaries, because such ventures benefit from the incentive alignment inherent to small, focused firms. Gromb and Scharfstein (2002) model whether a new venture should be pursued within the established firm ("intrapreneurship") or outside the firm. Their mechanism rests on the higher-powered incentives of the entrepreneur. When the new venture has potentially large payoffs and high failure risk, the benefits of locating the idea outside the firm in a new business outweigh the safety net benefits of intrapreneurship.

This mechanism also predicts that R&D-induced, employee-founded startups are more likely to develop new ideas or technologies that are far from the parent firm's core focus and have poor complementarities with its existing assets. When a firm rejects a new idea that would diversify the firm's activities, employee-founded startups may be a byproduct. Permitting employee entrepreneurship is likely to be most appealing when the lost ideas tend to be peripheral to the firm, rather than in its area of core focus. Relatedly, when assets are more complementary there is more potential for hold-up and thus benefits from locating the assets within a single firm (Hart and Moore 1990). Empirical work finds a negative correlation between firm performance and diversification (Lang and Stulz 1994, Schoar 2002). There is also practitioner evidence that sustained corporate success demands discipline in rejecting good opportunities that would make the firm's activities excessively diffuse (Collins 2009, McKeown 2020). Therefore, we expect that the firm will be most permissive towards employee entrepreneurship with businesses that would be diversifying if pursued in-house, and so expect that R&D-induced startups will more often be in different broad industries from their parents.

In sum, there may be benefits to developing a risky new idea in a new venture rather than within the parent firm, which could lead R&D to have a positive effect on employee entrepreneurship. In turn, this suggests that external capital markets, such as VC, will be better sources of financing for the idea than internal capital markets. Further, a permissive policy towards employee-founded startups could allow the firm to maintain the benefits of focusing on existing products and customers and could dynamically incentivize employees to maximize effort. Our hypotheses represent channels for R&D spillovers, since the R&D-induced startups are by-products of contracting frictions (Romer, 1990).

3 Data

We use data from five sources: Compustat, the U.S. Census Bureau's Longitudinal Business Database (LBD) and its Longitudinal Employer-Household Dynamics (LEHD) data, ThomsonOne VentureXpert, and the NBER Patent Data Project. This section describes these sources (Section 3.1), explains how we construct employee outcomes (Section 3.2), and presents summary statistics (Section 3.3).

3.1 Data sources

Our measure of corporate innovation investment is R&D expenditure as reported in 10-K filings and provided by Compustat. We restrict the sample of potential "parent" firms to public firms with positive R&D.[11] We primarily use log R&D but show that the results are robust to using R&D divided by total assets. We merge Compustat to the restricted-access LBD using a Census-provided crosswalk. The LBD is a panel dataset that tracks all U.S. business establishments with paid employees. An establishment is a discrete physical location operated by a firm with at least one paid employee. The LBD contains a unique firm-level identifier, *firmid*, which links establishments that are part of the same firm. Incorporated businesses rather than sole proprietorships or partnerships comprise about 83 percent of the LBD.[12] For further details about the LBD, see Jarmin and Miranda (2002). We use the 1978-2011 LBD for firm-level variables and to identify new firms. Following Haltiwanger et al. (2013), we define firm age using the oldest establishment that the firm owns in the first year the firm is observed in the LBD. A firm birth is defined when all of its establishments are new, preventing us from misclassifying an establishment that changes ownership as a startup.

A challenge when studying how R&D affects labor reallocation is that we must observe employees and track them from firm to firm. We solve this with the LEHD, which provides quarterly

[11] R&D expenditure is only available for public firms. We use firms with positive R&D for two reasons. First, firms that report R&D are likely qualitatively different from firms that do not in ways that might affect employee entrepreneurship, despite rigorous controls and fixed effects (Lerner and Seru 2022). Second, our primary specification will be focused on the intensive margin; since we use firm fixed effects, firms with zero R&D provide no variation. However, in a robustness check we include all Compustat firms and find similar results to the main specification.

[12] This is observable using the publicly available Census County Business Patterns data, which are built from the same Business Register that is the basis for the LBD.

firm-worker matched data. The data contain employees' wages, gender, race, place and date of birth, and citizenship status. The LEHD has been widely used in economic research (e.g., Tate and Yang 2015 and Goldin et al. 2017). In covered states, the LEHD includes over 96 percent of all private-sector jobs (BLS 1997, Abowd et al. 2009). About 12 percent of workers in year t are not in the LEHD in year $t + 3$. We document that this attrition rate is similar to those in the nationally representative Current Population Survey (CPS) in Appendix Section A.3, and our key independent variable—R&D—does not predict employee departures from the sample (Appendix Table A.1 Columns 5-6), suggesting that incomplete coverage likely does not affect our results.[13] Abowd et al. (2009) describe the construction of the LEHD data in detail.

Coverage begins in 1990 for several states and increases over time, ending in 2008. We have access to 31 states, in which we observe all employee-founded startups.[14] These states do not include the innovation hubs of California and Massachusetts. Their absence does not compromise the relevance and importance of our analysis. Our data include research-intensive states such as North Carolina and Texas; as a point of comparison, in 2017 venture capitalists invested about $1.8 billion in Texas and $1.9 billion in all five Nordic countries (Sweden, Norway, Finland, Iceland, and Denmark), which have individually been settings for important past research on entrepreneurship.[15] Public companies with significant R&D headquartered in our states include Amazon, Microsoft, ExxonMobil, General Dynamics, Red Hat, 3M, and Celgene, among many others. In our states, we also observe establishments of public firms headquartered in uncovered states. For example, IBM is headquartered in New York but has a research lab in Austin, Texas. Switzerland-headquartered pharmaceutical giant Novartis has a large R&D operation in New Jersey. In total, 42.4% of Compustat R&D expenditures are represented by the 31 states for which we have LEHD coverage.[16]

Our data are also representative beyond incumbent R&D. Sixty percent of U.S. employment,

[13]The CPS tracks workers for a maximum of 16 months. In the CPS data, among private sector employees who are observed 15 months later, about 9.9 percent drop out of the employment sample (data available here).

[14]The states we observe are: Arkansas, Delaware, Florida, Georgia, Colorado, Idaho, Iowa, Illinois, Indiana, Louisiana, Maine, Maryland, Missouri, Montana, Minnesota, New Jersey, New Mexico, North Carolina, Nevada, Oklahoma, Oregon, Pennsylvania, Rhode Island, South Carolina, Tennessee, Texas, Utah, Vermont, Virginia, Washington, and Wisconsin.

[15]See NVCA 2018 Yearbook here, and Pitchbook's European Venture report here.

[16]This assumes that a firm's R&D takes place entirely in its headquarter state. While some R&D may take place elsewhere, it is also true that firms in non-covered states likely perform R&D in covered states.

or about 150 million people, and 52 percent of U.S. inventors are located in our states.[17] Using LBD data, we calculate that 10.7 percent of firms are aged three years or less in our states, compared to 10.6 percent in all states. The sample is representative at the industry level, shown in Appendix Table A.2.[18] For example, in the 2002-08 period, Professional and Business Services (which includes information technology companies such as Google) represents 12.3 percent of employment in our sample states, and 12.8 percent of employment in other states. In sum, we observe meaningful quantities of innovation and high-growth entrepreneurship. A final note on this topic is that our research question may be especially policy-relevant to states that are not the leading innovation hubs. The effect of R&D on worker reallocation–especially R&D induced by local tax credits–is important for states actively seeking to promote innovation and entrepreneurship.

In the LEHD, workers are identified with firms' state reporting units, or State Employer Identification Numbers (SEINs). Each SEIN contains state and industry information. We link SEINs to firms in the LBD using federal employer identification numbers present in both datasets. For ease of exposition, we term SEINs "establishments."[19] We do the linkage in the first quarter of each year since the annual LBD measures employment and payroll in March. We drop SEINs with less than 10 employees, as they tend to have noisy reporting.[20] This yields an annual panel of public firm establishments (i.e., SEINs), in which employees are observed as of the first quarter of each year.

To identify VC-backed startups, we use the Puri and Zarutskie (2012) link from ThomsonOne VentureXpert to the Census Business Register. We use patent data from the NBER Patent Data Project, which includes patents linked to Compustat through 2006. We employ several annual, patent-based variables at both the firm and industry level. These are the number of patent classes a firm or industry

[17] We use LBD data to calculate the former statistics and comprehensive USPTO patent data between 1990 and 2005 to calculate the latter.

[18] Specifically, we compare our data to data from the Bureau of Labor Statistics (BLS) Current Employment Statistics Survey from 1990-2008. According to the BLS, employment data come from a voluntary, state level stratified sample of firms that is adjusted for population using monthly state unemployment insurance records. First, we divide state-industry employment by total state employment across all states for each year, and then take the average of this object across years. We conduct the same calculation for states out of our sample. A second calculation considers the share of people employed in an industry in our sample states versus the other states. The results are in Appendix Table A.3. The share of employment for each industry is quite similar to the overall share of employment we observe.

[19] We use SEINs for linking employees in the LEHD to firms in the LBD because this provides a more precise match than linking at a lower level of analysis in the LEHD data (e.g., SEINUNIT, which has links to employees that are imputed and do not reflect where an employee actually works).

[20] We obtain similar results if we drop establishments with less than five or less than fifteen employees.

patents in, the number of patents, the number of forward and backward citations, and patent generality. Generality is higher when forward citations are in many classes. All variables are defined in Appendix Table A.4.

3.2 Identifying employee outcomes

Our final sample consists of an annual panel of public firm establishments in 31 states between 1990–2005, with employees followed through 2008. We measure departure rates at the establishment level, rather than the firm level or employee level, for several reasons. First, public firms often have operations in multiple industries and/or states. Establishment-level analysis permits including as controls industry-year and state-year fixed effects as well as establishment workforce characteristics and wages. Note that both industry and state are defined at the establishment level. Second, the more disaggregated data allow cross-sectional tests. For example, Amazon has warehouses and business service offices. Using establishment-level data, we can test if the effect of R&D within Amazon is different in business offices than in warehouses. Third, we do not use employee-level observations in our primary approach as they create an artificially large amount of power, since variation in R&D occurs at the firm level. However, the results are robust to employee-level analysis (Section 5.4).

We begin by observing workers at public firm establishments (denoted e) in the first quarter of year t, and the quantity of R&D investment in year $t-1$. Using longitudinally consistent individual identifiers in the LEHD, we follow employees one, two, and three years after year t. They may move to other incumbent firms, which we define as firms that exist before year t, stay at the R&D-performing firm, or drop out of the employment sample. About 94 percent of workers in the "drop out" group go to unemployment or exit the labor force. Only around six percent depart to locations outside our LEHD data coverage.[21]

Identifying employee-founded startups is challenging because the LEHD does not provide information on equity ownership. Our aim is to capture members of new firms' founding teams, who contribute crucial early-stage ideas and skills. The measure is in line with prior research using

[21] We do not include departures to other establishments within the same firm as outcomes in any of the analyses.

executive teams, including Gompers et al. (2005). We focus on the highest earners to identify founders and early employees with important human capital. Azoulay et al. (2020) find that a firm's top three initial earners usually include the firm's owners. This is because the Unemployment Insurance earnings data that is the basis for the LEHD must be filed for all employees, including owners who actively manage the business and are required by law to pay themselves reasonable wage compensation.[22] Our primary definition of an employee-founded startup is a firm founded between t and $t + 3$ in which any of the parent firm establishment's employees at year t is among the top five earners as of $t + 3$, following Babina (2020).[23] We show the effect on a range of other entrepreneurship measures to demonstrate robustness (see Section 5.4). From the perspective of identifying entrepreneurs in the sense of a single initial owner-founder, our method likely yields some false positives. However, this is not a concern because our goal is to focus on a broader concept of entrepreneurship in the sense of the early, high skill employees of a new firm.

3.3 Summary statistics

Table 1 contains summary statistics at the parent firm-year, parent establishment-year, and employee-founded startup levels, respectively. We show the mean for indicator variables, as well as the quasi-median and the standard deviation for continuous variables.[24] Our dependent variables are measured at the establishment-year level as explained above. For all outcome variables, we divide by e's employment in year t to give, for example, the establishment rate of employee entrepreneurship. On average, 26 percent of an establishment's employees move to other incumbents (firms that existed before year t employment at a public firm), and 7.2 percent move to firms that are less than 10 years old. The rate of employee entrepreneurship is 1.3 percent. Similarly, Kerr et al. (2022) find in LBD/LEHD matched data that 1.7 percent of workers become entrepreneurs over a four-year period.

[22] See here.

[23] The lag is motivated by the time necessary to start a firm and to identify the effects of R&D, which might not be immediate. We examine the timing of departures in Section 5.4.

[24] As Census disclosure procedures prohibit the disclosure of percentile values, we approximate the median with a quasi-median, which is estimated using a 99 percent weight on observations within the interquartile range and a 1 percent weight on the remaining observations. The number of observations and all estimates in the tables are rounded according to the Census disclosure requirements.

The bottom set of variables describe the 108,000 employee-founded startups identified in the LBD. In their first year, the new firms have on average 12 employees, and 70 percent are incorporated businesses. Since we are the first to match VC investments data to US firm-worker matched data, we also document that two percent of the employee-founded startups ever receive venture capital funding, which is 18 times the rate of VC backing among the whole population of new employer firms (Puri and Zarutskie 2012).

There may be concern that employees depart the firm and found a startup in an entrepreneurial hub such as Silicon Valley. This may sometimes occur, but it is likely rare. Guzman (2019) finds that only 0.2 percent of new firms move to Silicon Valley in their first two years, based on data about all businesses registered in 26 states between 1988 and 2014. Note that if a startup has at least one employee in the home state before moving, we observe them. To the degree that former employees move to an uncovered state before hiring an employee, these individuals are counted in our outcome variables as having dropped out of coverage. If this phenomenon is important, we should observe a positive effect of R&D on dropping out of the labor force, but we do not.

4 Empirical approach

The primary estimation strategy, a tightly controlled fixed effects regression, is introduced in Section 4.1. In Section 4.2, we explain our instrumental variables strategy, and discuss market-level R&D in Section 4.3.

4.1 Main model

Our main approach consists of variants on the OLS regression in Equation 1, where e denotes an establishment, f a firm, and t the year:

$$\text{Employee outcome}_{e,f,t+3} = \beta \ln\left(\text{R\&D}_{f,t-1}\right) + \text{Firm FE}_f + \text{Industry-year FE}_{e,t} + \text{State-year FE}_{e,t} + \xi \text{Controls}_{f,t} + \zeta \text{Controls}_{e,t} + \varepsilon_{e,f,t}. \quad (1)$$

Our primary independent variable is log R&D among firms with positive R&D.[25] We use a flow R&D variable so that we can identify the effect of within-firm changes in R&D; using a constructed stock measure would make this approach empirically impractical as there would be much less year-to-year variation. We use firm fixed effects so that identifying variation comes from within-firm changes, thus controlling for time-invariant differences across firms. In this way, we mainly compare the same employees before and after an R&D change, so effects do not stem from firm-worker matching (we also find that the effect is not driven by recent hires). To use a concrete example, rather than comparing employee departures from Microsoft with those from Walmart, within-firm estimation allows us to compare employee departures within Microsoft between periods of higher and lower R&D.

We also control for industry-year fixed effects at the establishment level to control for changes in investment opportunities. We primarily use SIC three-digit codes but show robustness to four-digit codes. State-year fixed effects control for regional shocks, which may affect investment opportunities at incumbents as well as entrepreneurship. State is also defined at the establishment level. Note that these fixed effects ensure that proximity to omitted states should not be correlated with any mismeasurement. We cluster standard errors by firm, following the suggestion in Bertrand et al. (2004) to cluster by the cross-sectional unit that is the source of treatment variation.

Time-varying establishment and firm controls address other concerns. Our baseline controls are as follows. We include establishment number of employees in case, for example, smaller establishments have more focused or autonomous cultures and thus lead to more employee entrepreneurship. We also control for the establishment's average wage, defined as payroll divided by employment. We further include the following firm-level controls: return on assets, sales growth, Tobin's Q, asset tangibility (measured as plant, property and equipment investment divided by total

[25] The results are robust to including non-R&D performing firms and using R&D/Assets.

assets), size (log total assets), cash holdings, age, and diversification (an indicator for firms having establishments in multiple SIC three-digit industries).

In robustness tests, we also include three other sets of controls, all of which are measured in the year after R&D is observed. First, employee controls are establishment average age, female, white, foreign, education, tenure, and experience. Second, we control for patent activity: firm log patent classes, patents, forward citations, and backward citations. Third, we control for market R&D and market R&D squared. We define market R&D as the sum of all R&D expenditures of firms in the firm's establishment industry-state-year. Then we subtract the firm's own R&D, and finally divide by 1,000 so that the coefficients are not very small.

4.2 Instrument for R&D

Endogeneity may bias the OLS estimates of Equation 1. For example, an unobserved new technological opportunity may jointly engender parent R&D and employee reallocation. This could lead to an increase in market-level R&D that is in fact the driver of mobility, rather than origin-firm R&D specifically. The data do not actually support this concern, but we hold fixed destination-firm R&D by instrumenting for parent firm R&D using changes in the tax price of R&D, which were first implemented in Bloom et al. (2013) and are also used in König et al. (2019). This instrumentation approach rests on the assumption that the parent firms we study – all of which are relatively large publicly traded firms – have tax liabilities that make them sensitive to the tax price of R&D, which startups have little if any tax liability and also conduct R&D as part of an initial growth and survival strategy, and thus are not sensitive to changes in R&D tax credit generosity.

Specifically, we employ two instruments: one is the firm's state R&D tax price ($\rho^S_{f,t}$), and the other is based on the federal R&D tax credit ($\rho^F_{f,t}$). The former is firm-state-time specific and is calculated using inventor locations. The latter is firm-time specific because the definition of expenditure that can be applied to the federal R&D tax credit depends on a firm-specific "base." The Appendix contains details about the instruments' construction, which closely follows Bloom et al. (2013). The first stage regression is:

$$\ln(R\&D_{f,t}) = \beta_1 \ln\left(\rho^S_{f,t}\right) + \beta_2 \ln\left(\rho^F_{f,t}\right) + \text{Firm FE}_f + \text{Industry-year FE}_{e,t} \\ + \text{State-year FE}_{e,t} + \xi\text{Controls}_{f,t} + \zeta\text{Controls}_{e,t} + \varepsilon_{e,f,t}. \quad (2)$$

As in the Equation 1, the IV estimation includes state-year and industry-year fixed effects. These will absorb any aggregate effects. The results are in Appendix Table A.5.

4.3 Role of market-level R&D

One challenge facing the design in Equation 1 is that the results could reflect R&D at the destination firm rather than the origin firm. In other words, employee mobility from one firm to another could reflect an increase in R&D spending at the destination firm, which is mirrored by the origin firm because these firms might be both experiencing a rise in R&D. We address this in the OLS models by controlling for market R&D in the regression, as described above.

However, this concern is also one reason why our paper focuses on the outcome of employee entrepreneurship, rather than studies mobility more broadly. The IV is valid only for the outcome of employee entrepreneurship because startups do not typically benefit from R&D tax credits. We can therefore create a setting in which the destination firm–the employee-founded startup–has its R&D held fixed while the tax credit price of R&D moves the origin (i.e. parent) firm R&D expenditure. This allows us to test whether growth options created by R&D are often located in a new, standalone firm. As explained above, this could reflect various contracting channels between the firm and the employee; for example, allowing the employee to "walk away" with some ideas may be ex-ante optimal for attracting the best talent. Regardless, we are focused on testing the mobility of employees from large firms to startups that results from corporate investment in innovation. The decision to leave with an idea and pursue it in one's own enterprise requires holding destination firm R&D fixed, which we accomplish through the IV and, to a lesser extent, via the controls in the OLS model.

In a falsification check, we assess whether there are effects in the OLS model on departures to other incumbent firms. If our entrepreneurship result reflects market R&D and, through this channel, an increase at the destination firm, then we expect to also find a positive effect on departures to incumbents. While there are explanations for possible effects on other departures, which we describe in Section 5.4, if there are *no* effects, we can be much more confident that the main result reflects parent rather than destination startup R&D.

5 Results

This section first describes the main effect on employee entrepreneurship (Section 5.1), then studies how the effect varies with parent firm characteristics (Section 5.2), presents an IV model to address concerns about omitted variable bias (Section 5.3), and finally contains further robustness tests, including employee-level analysis and the effect on other types of labor mobility (Section 5.4).

5.1 Effect on employee entrepreneurship

Table 2 shows that parent firm R&D predicts a statistically significant increase in employee departures to entrepreneurship. The relationship between R&D and employee entrepreneurship is economically meaningful. In our most tightly controlled model (Column 5 in Panel 1), the coefficient of 0.109 implies that a 100 percent increase in parent firm R&D or approximately one standard deviation in R&D is associated with an 8.4 percent increase in the mean rate of employee departures to entrepreneurship, relative to the sample mean of 1.3 percent.[26]

The results are quite stable across specifications, consistent with the result not being driven by endogenous technological opportunities, which is the main concern in our setting. In Column 1, we use the simplest specification with firm and year fixed effects. In Column 2, we add the rich array of baseline controls (described in Section 4).In Appendix Table A.6 we report the coefficients of all the control variables.[27] While most controls have no predictive power for entrepreneurship, larger

[26]As R&D is in log units, the coefficient means that a 1 percent increase in R&D increases employee entrepreneurship by .109/100.

[27]Some controls are denoted with a lag ($t-1$) and others are not. This is because firm-level controls are measured when R&D is measured (year $t-1$), but establishment-level variables are measured when the employee snapshot is taken (first quarter of year t). We do not report them in further results because the Census Bureau strictly limits the

establishments have less employee entrepreneurship, consistent with prior work (Nanda and Sørensen 2010). The result continues to hold with industry and state fixed effects in Columns 3 and 4. Finally, in Column 5 we show that the effect is robust to including industry-year and state-year fixed effects, which should be correlated with technology shocks within industries and locations.[28] Panel 2 Columns 2, 3, and 4 show that employee characteristics, patent activity, and market R&D, all measured in the year after R&D is observed, do not attenuate the effect. Finally, Column 5 shows that including fine industry-state-year fixed effects does not affect the results.

We consider alternative measures of R&D in Appendix Table A.7. When the independent variable is an indicator for an above median change in R&D, the effect is .09, significant at the .01 level (Column 1). This implies that moving from the bottom to the top half of R&D changes increases the rate of employee entrepreneurship by seven percent. We find a similar effect on the number of new employee-founded startups (Column 2). This permits a back-of-the-envelope calculation that above-median relative to below-median R&D changes lead to 8,291 additional employee-founded startups over the sample period, which is 7.7 percent of all employee-founded startups in the data.[29] Appendix Table A.7 Columns 3 and 4 use indicators for high and low changes in firm R&D. Column 3 implies that moving from the bottom 90 to the top 10 percentiles increases the employee entrepreneurship rate by 12 percent. The effect turns negative for the bottom 10 percentiles of R&D change (Column 4). We also find that the effect is robust to using R&D divided by total assets, rather than the change in R&D (Column 5). This confirms that the effect is not an artifact of small changes in R&D.[30]

To shed light on employee entrepreneurship, we compare individuals who depart to entrepreneurship with those who depart to incumbent firms. We are particularly interested in wages at the parent firm as a measure of human capital. Appendix Table A.8 Panel A shows that individuals

number of coefficients we may disclose. The controls are at the firm level, except for employment and average wage which are at the establishment level. The only control with consistent predictive power is employment: employee entrepreneurship is negatively associated with the establishment's number of employees.

[28] It is also important to note that if a technology shock explained the relationship between R&D increases and employee entrepreneurship, then we would likely also observe increased departures to incumbent firms as well. However, we do not observe this in Table A.1.

[29] The calculation is as follows. As there are 329 employees in an establishment-year on average, the coefficient implies an increase of 0.23 employee-founded startups per establishment-year, which we multiply by the 36,000 establishment-years to arrive at 8,291 new firms.

[30] Another concern is that because some firms have multiple SEINs per state-year, our results could be driven by variation within firm-state-year that we are not capturing. Our effects are robust to excluding these firms.

moving to startup founding teams and to incumbents have similar age, tenure, and education. Future members of startup founding teams are somewhat more likely to be male and white. Founders' wages, however, are almost 50 percent higher. Appendix Table A.8 Panel B shows that the parent (R&D-performing) employing firm has similar characteristics across the two groups, and Panel C compares the destination firms. In sum, workers who depart to entrepreneurship have much higher wages than those who depart to incumbents, despite otherwise similar observable characteristics, suggesting that they are either managers, highly skilled, or both.

5.2 Parent heterogeneity

Entrepreneurial spillovers from R&D likely come from establishments close to the innovation process. Note that a new idea or technology need not leave the firm at its earliest stage; the firm may reject the new idea while it is in development or early commercialization. Therefore, R&D-induced employee entrepreneurs may emerge from various places in the firm. In general, we expect that R&D-generated ideas are more likely to be in high-tech establishments since high-tech industries are associated with technological spillovers. Entrepreneurship and industry are measured at the establishment level, and there is substantial within-firm variation in establishment industries.[31] Using an interaction between R&D and an indicator for the establishment being high-tech, we find that that the effect comes from high-tech establishments (Table 3 Column 1). There is no significant effect for non-high-tech establishments (the independent coefficient on R&D). The fact that high-tech establishments are responsible for the effect is consistent with startups being R&D spillovers.

Patenting activity provides a second source of confirmation. General-purpose patents are used in a wider array of fields (specifically, future cites are from a wider array of patent classes). We interact R&D with an indicator for the firm having above-median patent generality and find a significantly higher effect for these firms (Table 3 Column 2). Thus, consistent with the spillover interpretation, firms doing broader research have more employee-founded startups per dollar of R&D.

[31] The primary empirical design implicitly assumes that R&D is evenly distributed across establishments. Among firms in our sample, the quasi-median firm has establishments in five industries (measures using three-digit SIC codes). We define an establishment as "high tech" if its industry is in biotech, chemicals, software and business services, or high-tech manufacturing & R&D.

5.3 Instrumented effect on employee entrepreneurship

While an OLS strategy does not permit us to completely rule out omitted variable bias, the evidence in Section 5.1 does not support technological opportunities leading to both higher parent R&D and employee moves to startups. To further mitigate concerns about potential bias in the OLS estimates, we use the instrument from Bloom et al. (2013) based on changes in R&D tax credit policy. Table 4 contains the instrumented results using the same main specifications from Table 2. The coefficients are statistically significant and larger than the OLS results.[32] Our preferred specification, in Table 4 Column 3, is about five times the OLS estimate. The larger IV effect indicates that the subset of R&D expenditures affected by the tax credits leads to more employee entrepreneurship than the average increase in R&D.[33] This could reflect endogeneity biasing the OLS result downward. Alternatively, the local average treatment effect for compliers with the instrument may be larger than the population average treatment effect. As Angrist and Imbens (1995) and Jiang (2015) explain, this can lead to a larger IV effect even if the exclusion restriction is satisfied. Firms with R&D that is more sensitive to the tax price of R&D may have a higher causal effect of R&D on employee entrepreneurship.

There are two closely related explanations for such a phenomenon. The first is a correlation between propensity to generate employee-founded startups and adjustable R&D. Adjustable R&D may be more general or inventive, and thus more often yield innovations best suited to development outside the firm. It is not obvious why adjustable R&D would be more inventive, but we cannot rule it out. More plausibly, adjustable R&D is less crucial to the firm. The loss of the innovation output to employee-founded startups would then be less costly, implying lower ex-ante incentives to prevent employee entrepreneurship. That is, suppose the firm expects R&D to lead to some employee-founded startups. When the loss of these employees and ideas is expected to be costlier, the firm should increase R&D less in response to the tax price shock. The second and perhaps more straightforward explanation is that the marginal effect of R&D is higher than the average effect. OLS

[32]It may initially seem inconsistent that the state instrument uses patent locations to proxy for the location of R&D, yet patenting does not predict employee entrepreneurship (see Table 2 Panel 2 Column 3). The firms responsible for the IV result are patenting in general, but changes in their number of patents produced do not predict employee departures to entrepreneurship. It is also worth noting that the IV effect persists when using only the federal instrument.

[33]We do not use the IV estimator for interaction effects, such as in parent heterogeneity, because there is insufficient power to identify the interaction term.

estimates the effect of an additional dollar of average R&D. The IV strategy, which uses additional R&D tax subsidies to approximate increased R&D expenditure on the margin, better captures the effect on employee entrepreneurship of the "last" R&D dollar. The output from marginal R&D may be less costly to lose, perhaps because it is less predictable or farther from the firm's core focus.

5.4 Robustness tests

Alternative entrepreneurship measures To assess whether the result reflects the particular construction of the outcome variable, we use alternative measures of employee entrepreneurship in Appendix Table A.9. First, we consider the number of employee-founded startups rather than the number of departing employees. This is because team exits, where multiple employees depart together to a new firm, could explain the results. The dependent variable in Panel 1 Columns 1-2 is the number of employee-founded startups from an establishment, normalized by employment at $t = 0$. In Column 1, the coefficient implies that a 100 percent increase in R&D leads to a 5.8 percent increase relative to the mean, indicating that team exits do not explain the main results. Second, the effect is robust to including only incorporated employee-founded startups (Appendix Table A.9 Panel 1 Columns 3-4).

We also find a similar result using only the top three earners at the new firm rather than the top five (Appendix Table A.9 Panel 1 Columns 5-6). The result is further robust to restricting employee-founders to those employed at the new firm in the first year it appears in the LBD with positive employment (Panel 2 Columns 1-2). In Panel 2 Columns 3-6, we explore the timing of departures to entrepreneurship by looking at departure rates by year $t + 2$, as supposed to $t + 3$ in the main specification. In Panel 2 Columns 3-4, we consider only startups founded by year $t + 2$. We continue to find a positive, significant coefficient using this more immediate measure, although, as expected, the estimates are smaller than those for departures by year $t + 3$, which accumulates departures by $t + 2$ and $t + 3$. We then consider one-year-old startups in year $t + 2$,which captures more of a flow measure of departures. The effect of R&D remains significant (Panel 2 Columns 5-6). When we consider one-year-old startups in year $t + 3$, the effect is positive but insignificant (not reported). Therefore, R&D-induced departures to entrepreneurship occur in the first two years after the investment in R&D. Moreover, in a placebo test, we examine how R&D investments in year t predict *past* departure rates to entrepreneurship from in year $t-4$ to in year $t-1$, and find null results, further corroborating causal interpretation of the R&D effect on entrepreneurship (not reported).

Employee-level analysis We find similar results at the employee level. This suggests that neither by normalizing entrepreneurial departures by ex-ante establishment employment nor aggregating the data to the establishment level affect our results. Appendix Table A.10 shows OLS and IV effects of R&D on the probability an individual worker transitions to entrepreneurship, defined as being among the top five earners at a new firm within three years. The result in Column 1 implies that a 100 percent increase in R&D leads to a 5.3 percent increase in employee entrepreneurship at the worker level. (The mean of the dependent variable is 0.0091 percent.) As in our main estimates, the IV effect is larger (Columns 3 and 6). We use two sets of fixed effects, replacing state with state-year fixed effects in Columns 4-6. Throughout we include employee controls, specifically age, age squared, education, total experience in years, tenure at the firm, log earnings as of $t = 0$, and indicators for being female, white, foreign-born, and born in state. These address unobserved worker ability to the best degree possible given available variables in the LEHD. We do not use employee fixed effects because then the variation in R&D would come primarily from individuals moving between high- and low R&D-level-type firms. In that case, we would mainly capture the effect of selection of employees, which is an interesting question but is not the focus of this paper.

Reverse causality If the effect is causal, employee entrepreneurship should not predict R&D. To test this, we project R&D in year t on past employee entrepreneurship in Appendix Table A.11. In Column 1, we include one year of employee entrepreneurship, from year $t-2$ to year $t-1$. In Columns 2 and 3, we include two years ($t-3$ to $t-1$) and three years ($t-4$ to $t-1$), respectively. In all cases, the coefficient is insignificant. This additionally helps to allay the primary endogeneity concern, which is that an unobserved technological opportunity jointly causes R&D and employee entrepreneurship. Since the nature of a startup is to be adaptable and responsive to new opportunities, we expect startup founding to respond to such an unobserved new opportunity faster than corporate R&D. In contrast, we find that employee entrepreneurship occurs after R&D.

New hires Another possible source of endogeneity is that R&D may lead the firm to hire new employees, who are inherently more likely to start their own ventures than the average worker. In this case, workers with relatively short tenures would drive the effect. In fact, our effect is not driven by employees with short tenure. Further, R&D has no effect on entrepreneurship among employees hired within a year of when R&D is measured, and a positive effect (in both OLS and IV) among employees who were at the firm for at least three years before R&D is measured.[34] Therefore, hiring related to the increase in R&D does not drive the effect.

Effects on other employee outcomes We next examine the effects of R&D on employee retention, departures from the labor force, and moves to other incumbents. These outcomes help us to test the mechanism behind our main results and are also inherently interesting to study.

First, R&D investment could create new skills or ideas that improve employees' outside options at other incumbent firms (Herkenhoff et al., 2018). More generally, as explained in Section 4.3, if the effect of R&D on employee entrepreneurship reflects something about changes in market-level R&D, we would also expect to find an effect on moves to incumbent firms. Such an effect could also reflect other firms stealing ideas that the firm would prefer to retain in-house, talent poaching, or it may reflect the firm permitting employees to leave with ideas that it does not find especially valuable. Existing work on labor mobility and knowledge diffusion motivates this hypothesis, but to our knowledge, there are no existing tests of whether R&D investment predicts changes in labor mobility.

The effects of R&D on mobility to other incumbents is reported in Table A.1 using both the OLS and IV approaches. Columns 1 and 2 show that there is also no relationship with moves to other incumbent firms. The OLS coefficient in Column 5 permits us to rule out effects outside the 95 percent confidence interval bounds of -6.5 percent and 2.8 percent of the mean. This null result suggests that either R&D does not increase employees' outside options at peer firms, or that the parent prevents such departures (e.g., though labor contracts). In unreported results, we also find no relationships between patenting intensity and retention, departures from the labor force, or mobility to other incumbent firms,

[34] These results are unreported due to disclosure limitations.

suggesting that neither innovation inputs (R&D) nor outputs (patents) are related to non-entrepreneurial labor reallocation.

Second, R&D might increase employee retention if it generates growth, creating more internal opportunities and perhaps making the firm a more interesting place to work (Rosen 1986). Table A.1 Columns 5 and 6 show that there is no association between changes to firm R&D and the share of an establishment's employees who remain at the firm. The effect in Column 3 of -1.104 implies a decrease of 2.4 percent relative to the mean of 48 percent. The 95 percent confidence interval is between -5.3 percent and 0.01 percent.

Third, R&D investment may lead to automation or other structural changes that make labor redundant or skills obsolete, causing layoffs (Milgrom and Roberts, 1990; Acemoglu and Restrepo, 2018). Third, R&D may lead to automation or other structural changes that make labor redundant or skills obsolete, causing layoffs (Brynjolfsson and McAfee 2014). An alternative mechanism for R&D to increase layoffs is if it is associated with restructuring. Table A.1 Columns 5-6 show that there is no relationship between R&D and employee exits from employment, with the 95 percent confidence interval bounds in the OLS model being -2 percent and 2.1 percent of the mean. This implies that R&D does not lead to automation or other structural changes that make labor redundant. Together, these results show that R&D has no measurable effect on labor reallocation, helping to confirm that the effects on employee entrepreneurship reflect parent R&D specifically, and support our hypothesis that this occurs because growth options are reallocated via employees to startups.

Local labor markets We also ensured that our result does not reflect shocks to the local labor market. One way we do this is through the control for industry-state-year fixed effects in Table 2. We also show that all the main results are robust to MSA-by-year fixed effects in Appendix Tables A.12 and A.13. These fixed effects control for variation in R&D between labor markets, under the assumption that workers switch mostly within and not between these markets. Specifically, in Appendix Table A.12 Column 1, we repeat the main result from Table 2 with the new controls, and see a similar result of 0.103 relative to the same specification without these MSA-by-year fixed effect in our main table of 0.106 (Column 3 of Table 2). The remainder of Appendix Table A.12 replicates Table A.7, and

Appendix Table A.13 replicates Table 5.[35]

6 Mechanisms

In Section 2, we hypothesized that incumbent firms may permit employees to leave and found new firms with some of the ideas that serendipitously emerge from R&D. This section considers evidence for two predictions that emerge from this mechanism: (1) R&D-induced employee-founded startups will be more likely to be risky and potentially high-growth (Section 6.1); and (2) R&D-induced, employee-founded startups are more likely to develop new ideas or technologies that are far from the parent firm's core focus and have poor complementarities with its existing assets (Section 6.2). Next, we discuss the role of incomplete contracting (Section 6.3). Finally, we present evidence against alternative mechanisms in Section 6.4.

6.1 High-risk, high-growth

We examine whether parent R&D is associated with high-risk, high-growth startup characteristics in Table 5 Panel 1. Here, analysis is conducted at the startup level and the sample consists of all employee-founded startups in our data.[36] Our first test concerns VC backing because VC-backed startups are widely known to be risky, associated with new-to-the-world ideas, and potentially high growth (Kaplan and Lerner 2010, Gornall and Strebulaev 2021)—the type of startups that are an important source of spillovers. While VC-backed startups comprise just 0.11% of all US firms, they account for 5.3%–7.3% of employment in the US (Puri and Zarutskie, 2012). The dependent variable in Column 1 is one if the employee-founded startup receives VC. The coefficient on parent firm R&D is 0.007, which is significant at the .01 level. This implies that a 100 percent increase in R&D predicts a 35 percent increase in the chances that an employee-founded startup is VC-backed.

[35] We do not include the other tables with these controls for parsimony, but the results are all robust.

[36] These regressions do not include firm fixed effects. As a relatively rare event, entrepreneurship provides limited variation for within-firm comparisons. The regressions should, therefore, be interpreted as well controlled associations. To the degree the results have pointed toward a causal relationship, this is appropriate for exploring mechanisms.

Levine and Rubinstein (2017) show that incorporation is a good indicator of high-growth intent in the sense of "business owners engaged in non-routine, innovative activities." Consistent with this, we find that R&D-induced startups are more likely to be incorporated (Column 2). If R&D leads to the diffusion of new technologies, we also expect high-risk, high-growth ventures emerging from R&D to be high-tech, which is the case (Column 3). Further, R&D induces employee-founded startups with higher wages than the average employee-founded startup, suggesting that they employ higher skilled labor, which is more likely to be a channel of spillovers (Column 4). Finally, we consider the rate of exit, which we view as a proxy for risk, comprised primarily of firm failures but likely includes a small share of acquisitions. In Column 5, the dependent variable is one if the startup exits within five years (starting from year $t+3$, where t is the year in which we measure R&D). We find a positive, significant association with R&D. In sum, relative to the average employee-founded startup, those induced by R&D are more likely to be high-risk and high-growth.

6.2 Costly diversification

The costly diversification mechanism fits well with the interpretation of the IV results in which ideas leading to employee entrepreneurship are more likely to come from the last dollar of R&D than the first. In this light, the IV strategy isolates the mechanism: marginal R&D more often generates ideas far from the firm's core focus, some of which spill into employee-founded startups. The following subsections consider cross-sectional and supply chain evidence.

6.2.1 Cross-sectional evidence

We begin by comparing parent and startup industries. In Column 6 of Table 5 Panel 1, the dependent variable is one if the employee-founded startup is in the same two-digit SIC classification as its parent. Two-digit industries are quite broad; examples are Business Services, Health Services, and Coal Mining. We find that a 100 percent increase in R&D makes it 4.2 percent less likely that the employee-founded startup is in the same industry as its parent. [37]

[37] While SIC industries might be coarse, unfortunately, we cannot use more granular measures such as the Hoberg-Phillips industry (Hoberg and Phillips 2016) because this is measured from financial disclosures that do not exist

It may initially seem counterintuitive that R&D leads employees to found firms in different industries. However, let us return to the two examples from the Introduction. First, in 1894, Henry Ford left Thomas Edison's Illuminating Company to start his own vehicle manufacturing business. Edison would be in SIC 49 (Electric, Gas and Sanitary Services), while Ford is in SIC 37 (Transportation Equipment). Yet Ford relied on mechanical and electrical engineering advances made at Edison's company. Second, in 1999, Michael Rosenfelt left Micron Electronics to found Powered Inc., an online education company. Micron Technology is in SIC 36 (Electronic and other Electrical Equipment), while Powered, Inc. would be in either SIC 73 (Business Services), or SIC 82 (Educational Services). In these examples, an R&D-intensive parent spawned a new firm in a different two-digit SIC sector, but the underlying idea or skill was related to the parent's intellectual capital. These examples highlight how SIC assignments reflect the firm's product market more than its technology. It seems likely that R&D-induced startups employ innovation related to the parent's technology but apply them to a different market, which is consistent with technological spillovers.

6.2.2 Supply chain relationships

To explore links between the startups and their parents, we consider supply chain relationships. We use the U.S. BEA annual input-output tables to create annual measures of supply chain closeness between the parent firm's industry and the startup's industry. The measures assign one party to be upstream and the other to be downstream. The first measure is "downstream closeness," which is the downstream industry's share of the upstream industry's product. The second measure is "upstream closeness," which is the upstream industry's share of what the downstream industry uses.[38] For both measures, a higher value means they are closer.

The results are in Table 5 Panel 2. We first assign the parent to the upstream industry, and

among new firms.

[38]Downstream closeness is built using the BEA "Make table," which contains the production of commodities by industries, where industries are in rows, and the Columns represent commodities (products) that the industries produce. Given industry pair A and B, if A is the "industry" and B is the "commodity," downstream closeness is B's share of A's row. Upstream closeness is built using the BEA "Use table," which contains the use of commodities by intermediate and final users, where commodities are in rows, and the columns represent the industries that use them. Given industry pair A and B, if A is the "industry" and B is the "commodity," upstream closeness is B's share of A's column. We use two-digit NAICS codes. Data are available here.

the employee-founded startup to the downstream industry. There is a positive association with the "downstream closeness" measure (Column 1). This means that R&D-induced startups tend to buy a relatively larger share of the parent's product than the average employee-founded startup.[39] The effect of "upstream closeness" is negative, which means that the parent's product tends to make up a relatively smaller share of the R&D-induced startups' inputs (Column 2). Therefore, R&D-induced startups tend to be downstream from the parents but require a broad array of inputs–not just from the parent, but from other industries as well. When we assign the employee-founded startups to the upstream industry, and the parent to the downstream industry, we find no effect of downstream closeness (Column 3). We find a weak positive effect of upstream closeness (Column 4), implying that the R&D-induced startup's product tends to make up a somewhat larger share of the parent's inputs.

These results demonstrate a tie between R&D-induced startups and their parents. However, the R&D-induced startup departs from the parent in that it requires more inputs from other industries. With diverse required inputs, many of the transactions required for commercialization would be outside the parent firm anyway, helping to explain why vertical integration might not be optimal. This is consistent with the R&D-generated new venture being farther from the parent's core focus. The fact that the parent's product makes up a relatively smaller share of the R&D-induced startups' inputs also offers one pathway for spillovers: By purchasing a broader array of goods, R&D-induced startups connect to new supply chains and more sectors, facilitating broader knowledge diffusion.

6.3 Incomplete contracting

R&D investment yields innovations in a highly uncertain, serendipitous manner. Sometimes, the outputs will not be useful to the firm. One indicator of this is if the effect of R&D on employee entrepreneurship emerges from those innovation outputs over which the firm does not establish explicit, contractible ownership (Kim and Marschke 2005). Patents measure R&D outputs that the firm has chosen to appropriate. We find that neither the number of patents nor the number of patent

[39] To the degree the spawn purchases from the parent, this does not imply that the parent benefits from the spawn. For example, if both industries are competitive, the spawn can presumably buy the input from alternative suppliers at the market price.

citations have an effect on employee entrepreneurship (Table 2 Panel 2 Column 3).

To explore whether the employee-founded startups and parents are in sectors that tend to share knowledge, we create two measures of patent citation flows between industries. These are derived from citation flows at the patent class-year level. One is from startup industry to parent industry (inflows), and the other is from parent industry to startup industry (outflows).[40] After constructing these measures at the class-year level, we assign patent classes to industries using the patent-to-SIC concordance developed by Kerr (2008). We can then estimate the effect of R&D on the chance that the employee-founded startup is in the top 5% of the knowledge sharing distribution. The estimates are reported in Appendix Table A.14, and are at at the employee-founded startup level. We calculate the measures both including self-citations (odd columns) and excluding them (even columns). In all four models, we find zero effects. This supports the conclusion that our results reflect R&D output that is not patented. Ellison et al. (2010), who also use this knowledge-sharing measure, find weak effects, and suggest that "knowledge sharing... may be captured more by input-output relationships than by these citations." We view these null results for contractible outputs (patents) as important evidence of the role of incomplete contracting in innovation. Theoretically, it is natural that innovation spillovers–those R&D outputs that cross the firm boundary–are primarily composed of non-contractible outputs.

6.4 Alternative mechanisms

The cross-sectional evidence presented above does not rule out alternative mechanisms. This section considers four additional channels.

6.4.1 Project management skills

Exposure to R&D could make employees more productive as entrepreneurs if they gain experience managing new projects. This channel may play a role, but three pieces of cross-sectional evidence

[40] Specifically, for patent classes A and B, inflows are B's cites of A as a share of the total cites to A. Outflows are A's cites of B as a share of all the citations from A. We are especially grateful to Bill Kerr for his help with this exercise.

suggest that it is unlikely to be the primary driver. First, we expect that capital expenditure would have a similar effect on employee entrepreneurship if the channel were skills, because it is likely to create project management skills. Instead, Appendix Table A.6 shows that there is no effect of total investment or PPE investment on employee entrepreneurship. We would also expect R&D-induced startups to come from small parents. This is because small firm employees tend to have a broader scope of work (Stuart and Ding 2006, Sørensen 2007). Instead, large firms drive the effect (Table 3 Column 3). Third, we expect that there is more opportunity for entrepreneurial learning at young firms. However, we find no effect of an interaction between R&D and firm age (unreported).

A related mechanism is whether firms that have recently gone public drive the effect. In this case, it may reflect employees "cashing out" their stock options rather than R&D (Babina et al. 2020). In Table 3 Column 4 we interact R&D with an indicator for having had an IPO within the last three years. The interaction is positive, but it is insignificant and does not attenuate the main effect.

6.4.2 Idea stealing

Another possibility is that employees "steal" ideas from their employer. Several pieces of evidence suggest that idea stealing is not the main mechanism explaining the effect of R&D on employee entrepreneurship. First, we expect a stealing mechanism to attenuate in states that enforce noncompete covenants. Noncompetes restrict employees from working for a competing firm within the state for a specified period of time. It has been found that noncompete enforcement reduces local R&D spillovers (Belenzon and Schankerman 2013) and within-state inventor mobility (Marx et al. 2015). The main result persists in states that enforce noncompetes, and there is no significant effect on an interaction between R&D and an indicator for being in a weak enforcement state (unreported). Second, any effect should attenuate when intellectual property is easier to protect (this also makes it is easier to contract on innovation effort). We do not find that the effect varies with industry patentability. Finally, there is a revealed preference argument. By virtue of observing the robust phenomenon of R&D-induced employee entrepreneurship, the parent either chose not to develop the idea in-house or chose not to take steps to prevent the employee-founded startup. Such steps could include increasing the employee's compensation to retain him, or not conducting the R&D at all.

6.4.3 Employee involvement with the change in R&D

There is concern that the employee who departs for entrepreneurship causes the R&D increase or is hired as a result of it. The first possibility is obviated by the IV strategy, where we identify the effect of R&D on employee-founded startups using only variation in R&D explained by its tax price, which the employee does not control. The second possibility is unlikely because we find a significant result using only workers with above-median tenure, as discussed earlier.

6.4.4 Internalization of startup benefits

It may be that the parent captures some of the startup's benefits, perhaps through a licensing or investment contract.[41] If the parent wholly owns the spinoff and captures all its benefits, then the effect we observe is not an R&D spillover in the sense of being a benefit of R&D that accrues to a firm besides the R&D-performing firm. The data do not support full internalization. First, we expect parent-supported spinoffs to start at a larger scale than a typical bootstrapped startup. We find no relation between initial employee-founded startup size and parent R&D (unreported). Second, spinoffs or parent reorganization should sometimes maintain the same establishment. Startups are defined in our data as firms with no prior activity at any of their establishments.

We also look for internalization in an out-of-sample test based on the underlying data in Gompers et al. (2005). This exercise is described in detail in Appendix Section A.2. We examine what share of the 6,499 unique VC-backed startups in the Gompers et al. (2005) data was acquired by startup executives' previous employers. This should yield an upper bound on internalization. Just 2.3 percent of the 9,152 unique parents match to an investor or acquirer, providing evidence that parents rarely invest in or acquire employee-founded startups. Consistent with the out-of-sample test, there is no effect of an interaction between R&D and the parent having a corporate VC program.

[41] An alternative is that patents jointly held or otherwise licensed across the parent and the startup permit a degree of internalization. Since patenting has no effect on employee entrepreneurship, this seems unlikely.

7 Conclusion

The outcomes of innovation investment are uncertain, serendipitous, and difficult to contract on. Employees, with their inalienable and portable human capital, create a porousness to the firm's boundary, providing an avenue for R&D outputs to leak to other firms. Corporate research effort yields ideas that are first embodied in people and then ultimately in new types of capital inputs (Mankiw et al. 1992, Jones 2002). In this way, R&D imparts new skills and ideas to employees. This paper shows that some of these outputs are reallocated to startups through employee mobility. We extend the literature on innovation spillovers by demonstrating a real effect of corporate R&D investment: new firm creation. Our evidence is consistent with high-tech startups being a new channel for R&D spillovers. There are private spillovers to the entrepreneur and other equity holders, and social value from new jobs created or the commercialization of new ideas.

Existing literature has emphasized how by generating monopolistic rents, incumbent R&D may stifle new firm creation (Bankman and Gilson 1999, Acemoglu et al. 2013). Our results offer a contrasting perspective and have implications for policy: The effect of R&D on employee entrepreneurship implies greater corporate underinvestment in R&D relative to the social optimum than previously thought. Also in contrast with a common assumption, we find no relationship between R&D and labor mobility to other incumbents. Finally, much of the innovation literature focuses on innovation outputs, especially patents and patent citations. We show that there is no predictive power of patents on any kind of labor mobility.

Human capital is central to modern theories of economic growth (Romer 1990, Jones 2014). Though the literature has focused on schooling (Card 2001, Cunha and Heckman 2007), firms also play a role in human capital formation. We document a likely unintended consequence of innovation inputs: employee-founded startups. Consistent with influential theories of the firm, these R&D-induced startups are more likely to be high-risk and potentially high-growth. They seem to reflect projects rejected by the firm because they are far from existing activities. With tight incentive alignment between owners and managers, startups present an attractive venue for these projects.

References

Abowd, John M, Bryce E Stephens, Lars Vilhuber, Fredrik Andersson, Kevin L McKinney, Marc Roemer, and Simon Woodcock. 2009. The LEHD infrastructure files and the creation of the Quarterly Workforce Indicators. In *Producer dynamics: New evidence from micro data*. University of Chicago Press, 149–230.

Acemoglu, Daron, Ufuk Akcigit, Nicholas Bloom, and William Kerr. 2013. Innovation, reallocation and growth. *NBER Working Paper No. 18993* .

Acemoglu, Daron and Pascual Restrepo. 2018. The race between man and machine: Implications of technology for growth, factor shares, and employment. *American Economic Review* 108 (6):1488–1542.

Agarwal, Rajshree and Sonali K Shah. 2014. Knowledge sources of entrepreneurship: Firm formation by academic, user and employee innovators. *Research Policy* 43 (7):1109–1133.

Aghion, Philippe and Xavier Jaravel. 2015. Knowledge spillovers, innovation and growth. *The Economic Journal* 125 (583):533–573.

Aghion, Philippe and Jean Tirole. 1994. The management of innovation. *The Quarterly Journal of Economics* 109 (4):1185–1209.

Akcigit, Ufuk and Sina T Ates. 2019. Ten facts on declining business dynamism and lessons from endogenous growth theory. Tech. rep., National Bureau of Economic Research.

Akcigit, Ufuk, Murat Alp Celik, and Jeremy Greenwood. 2016. Buy, keep, or sell: Economic growth and the market for ideas. *Econometrica* 84 (3):943–984.

Angrist, Joshua D and Guido W Imbens. 1995. Two-stage least squares estimation of average causal effects in models with variable treatment intensity. *Journal of the American statistical Association* 90 (430):431–442.

Atalay, Enghin, Ali Hortaçsu, Mary Jialin Li, and Chad Syverson. 2019. How wide is the firm border? *The Quarterly Journal of Economics* 134 (4):1845–1882.

Atkeson, Andrew and Ariel Burstein. 2019. Aggregate implications of innovation policy. *Journal of Political Economy* 127 (6):2625–2683.

Atkeson, Andrew and Ariel Tomas Burstein. 2010. Innovation, firm dynamics, and international trade. *Journal of political economy* 118 (3):433–484.

Audretsch, David B and Maryann P Feldman. 2004. Knowledge spillovers and the geography of innovation. In *Handbook of regional and urban economics*, vol. 4. Elsevier, 2713–2739.

Azoulay, Pierre, Benjamin F Jones, J Daniel Kim, and Javier Miranda. 2020. Age and high-growth entrepreneurship. *American Economic Review: Insights* 2 (1):65–82.

Babina, Tania. 2020. Destructive creation at work: How financial distress spurs entrepreneurship. *The review of financial studies* 33 (9):4061–4101.

Babina, Tania, Asaf Bernstein, and Filippo Mezzanotti. 2022a. Financial disruptions and the organization of innovation: Evidence from the great depression. *The Review of Financial Studies* Forthcoming.

Babina, Tania, Alex Xi He, Sabrina T Howell, Elisabeth Ruth Perlman, and Joseph Staudt. 2022b. Cutting the innovation engine: How federal funding shocks affect university patenting. Tech. rep.

Babina, Tania, Paige Ouimet, and Rebecca Zarutskie. 2020. Ipos, human capital, and labor reallocation. *Available at SSRN 2692845* .

Bai, John, Daniel Carvalho, and Gordon M Phillips. 2018. The impact of bank credit on labor reallocation and aggregate industry productivity. *The Journal of Finance* 73 (6):2787–2836.

Balsvik, Ragnhild. 2011. Is labor mobility a channel for spillovers from multinationals? Evidence from Norwegian manufacturing. *The review of economics and statistics* 93 (1):285–297.

Bankman, Joseph and Ronald J. Gilson. 1999. Why start-ups? *Stanford Law Review* 51:289, 293–95.

Baumol, William J. 1990. Entrepreneurship: Productive, unproductive, and destructive. *Journal of Political Economy* 98 (5, Part 1):893–921.

Belenzon, Sharon and Mark Schankerman. 2013. Spreading the word: Geography, policy, and knowledge spillovers. *Review of Economics and Statistics* 95 (3):884–903.

Bertrand, Marianne, Esther Duflo, and Sendhil Mullainathan. 2004. How much should we trust differences-in-differences estimates? *The Quarterly journal of economics* 119 (1):249–275.

Bilir, L Kamran and Eduardo Morales. 2020. Innovation in the global firm. *Journal of Political Economy* 128 (4):000–000.

Bloom, Nicholas, Mark Schankerman, and John Van Reenen. 2013. Identifying technology spillovers and product market rivalry. *Econometrica* 81 (4):1347–1393.

BLS. 1997. *Bls handbook of methods*. Bureau of Labor Statistics, Washington DC. U.S. Bureau of Labor Statistics, Division of Information Services, Washington D.C.

Borjas, George J. and Kirk B. Doran. 2015. Which peers matter? the relative impacts of collaborators, colleagues, and competitors. *Review of economics and statistics* 97 (5):1104–1117.

Brynjolfsson, Erik and Andrew McAfee. 2014. *The second machine age: Work, progress, and prosperity in a time of brilliant technologies*. WW Norton & Company.

Card, David. 2001. Estimating the return to schooling: Progress on some persistent econometric problems. *Econometrica* 69 (5):1127–1160.

Collins, James Charles. 2009. *How the mighty fall: And why some companies never give in*. Random House.

Combes, Pierre-Philippe and Gilles Duranton. 2006. Labour pooling, labour poaching, and spatial clustering. *Regional Science and Urban Economics* 36 (1):1–28.

Cunha, Flavio and James Heckman. 2007. The technology of skill formation. *American Economic Review* 97 (2):31–47.

Curtis, E Mark and Ryan Decker. 2018. Entrepreneurship and state taxation. Tech. rep., Board of Governors of the Federal Reserve System (US).

Decker, Ryan, John Haltiwanger, Ron Jarmin, and Javier Miranda. 2014. The role of entrepreneurship in us job creation and economic dynamism. *The Journal of Economic Perspectives* 28 (3):3–24.

Decker, Ryan A, John Haltiwanger, Ron S Jarmin, and Javier Miranda. 2016. Declining business dynamism: What we know and the way forward. *American Economic Review* 106 (5):203–07.

Ellison, Glenn, Edward L Glaeser, and William R Kerr. 2010. What causes industry agglomeration? Evidence from coagglomeration patterns. *The American Economic Review* 100 (3):1195–1213.

Foster, Lucia, John Haltiwanger, and Chad Syverson. 2008. Reallocation, firm turnover, and efficiency: Selection on productivity or profitability? *American Economic Review* 98 (1):394–425.

Frésard, Laurent, Gerard Hoberg, and Gordon M Phillips. 2020. Innovation activities and integration through vertical acquisitions. *The Review of Financial Studies* 33 (7):2937–2976.

Ganguli, Ina, Jeffrey Lin, and Nicholas Reynolds. 2020. The paper trail of knowledge spillovers: evidence from patent interferences. *American Economic Journal: Applied Economics* 12 (2):278–302.

Garcia-Macia, Daniel, Chang-Tai Hsieh, and Peter J Klenow. 2019. How destructive is innovation? *Econometrica* 87 (5):1507–1541.

Gennaioli, Nicola, Rafael La Porta, Florencio Lopez-de Silanes, and Andrei Shleifer. 2012. Human capital and regional development. *The Quarterly Journal of Economics* 128 (1):105–164.

Glaeser, Edward. 2011. *Triumph of the city: How urban spaces make us human*. Penguin Press. P. 47.

Glaeser, Edward L, Sari Pekkala Kerr, and William R Kerr. 2015. Entrepreneurship and urban growth: An empirical assessment with historical mines. *Review of Economics and Statistics* 97 (2):498–520.

Goldfarb, Brent, Gerald Marschke, and Amy Smith. 2009. Scholarship and inventive activity in the university: complements or substitutes? *Economics of Innovation and New Technology* 18 (8):743–756.

Goldin, Claudia, Sari Pekkala Kerr, Claudia Olivetti, Erling Barth et al. 2017. The expanding gender earnings gap: evidence from the lehd-2000 census. *American Economic Review* 107 (5):110–14.

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

Gornall, Will and Ilya A Strebulaev. 2021. The economic impact of venture capital: Evidence from public companies. *Available at SSRN 2681841* .

Griffith, Rachel, Rupert Harrison, and John Van Reenen. 2006. How special is the special relationship? Using the impact of US R&D spillovers on UK firms as a test of technology sourcing. *American Economic Review* 96 (5):1859–1875.

Griliches, Zvi. 1992. The search for R&D spillovers. *Scandinavian Journal of Economics* 94 (3).

Gromb, Denis and David Scharfstein. 2002. Entrepreneurship in equilibrium. Tech. rep., National bureau of economic research.

Grossman, Sanford J. and Oliver D. Hart. 1986. The costs and benefits of ownership: A theory of vertical and lateral integration. *Journal of Political Economy* 94 (691):693–95.

Guzman, Jorge. 2019. Go west young firm: agglomeration and embeddedness in startup migrations to silicon valley. *Columbia Business School Research Paper* (18-49).

Guzman, Jorge and Scott Stern. 2017. *Nowcasting and placecasting entrepreneurial quality and performance*, vol. 75. University of Chicago Press.

Haltiwanger, John, Ron S Jarmin, and Javier Miranda. 2013. Who creates jobs? small versus large versus young. *Review of Economics and Statistics* 95 (2):347–361.

Hart, Oliver and John Moore. 1990. Property rights and the nature of the firm. *Journal of Political Economy* 98:1152–53.

Heggedal, Tom-Reiel, Espen R Moen, and Edgar Preugschat. 2017. Productivity spillovers through labor mobility in search equilibrium. *Journal of Economic Theory* 169:551–602.

Herkenhoff, Kyle, Jeremy Lise, Guido Menzio, and Gordon Phillips. 2018. Worker mobility and the diffusion of knowledge .

Hoberg, Gerard and Gordon Phillips. 2016. Text-based network industries and endogenous product differentiation. *Journal of Political Economy* 124 (5):1423–1465.

Jaffe, Adam B, Manuel Trajtenberg, and Rebecca Henderson. 1993. Geographic localization of knowledge spillovers as evidenced by patent citations. *the Quarterly journal of Economics* :577–598.

Jarmin, Ron S and Javier Miranda. 2002. The longitudinal business database. *Available at SSRN 2128793* .

Jiang, Wei. 2015. Have instrumental variables brought us closer to the truth. *The Review of Corporate Finance Studies* .

Jones, Benjamin F. 2014. The human capital stock: a generalized approach. *American Economic Review* 104 (11):3752–77.

Jones, Charles I. 2002. Sources of us economic growth in a world of ideas. *American Economic Review* 92 (1):220–239.

Jones, Charles I and John C Williams. 1998. Measuring the social return to R&D. *The Quarterly Journal of Economics* :1119–1135.

Kaiser, Ulrich, Hans Christian Kongsted, and Thomas Rønde. 2015. Does the mobility of r&d labor increase innovation? *Journal of Economic Behavior & Organization* 110:91–105.

Kantor, Shawn and Alexander Whalley. 2014. Knowledge spillovers from research universities: evidence from endowment value shocks. *Review of Economics and Statistics* 96 (1):171–188.

Kaplan, Steven N and Josh Lerner. 2010. It ain't broke: The past, present, and future of venture capital. *Journal of Applied Corporate Finance* 22 (2):36–47.

Kerr, Sari Pekkala and William R Kerr. 2017. Immigrant entrepreneurship. *Measuring Entrepreneurial Businesses: Current Knowledge and Challenges* :187.

Kerr, Sari Pekkala, William R Kerr, and Ramana Nanda. 2022. House prices, home equity and entrepreneurship: Evidence from us census micro data. *Journal of Monetary Economics* 130:103–119.

Kerr, William R. 2008. Ethnic scientific communities and international technology diffusion. *The Review of Economics and Statistics* 90 (3):518–537.

Kerr, William R and Scott Duke Kominers. 2015. Agglomerative forces and cluster shapes. *Review of Economics and Statistics* 97 (4):877–899.

Kerr, William R and William F Lincoln. 2010. The supply side of innovation: H-1b visa reforms and us ethnic invention. *Journal of Labor Economics* 28 (3):473–508.

Kim, Jinyoung and Gerald Marschke. 2005. Labor mobility of scientists, technological diffusion, and the firm's patenting decision. *RAND Journal of Economics* :298–317.

Klepper, Steven. 2001. Employee startups in high-tech industries. *Industrial and corporate change* 10 (3):639–674.

Kondo, Jiro, Danielle Li, and Dimitris Papanikolaou. 2021. Trust, collaboration, and economic growth. *Management Science* 67 (3):1825–1850.

König, Michael D, Xiaodong Liu, and Yves Zenou. 2019. R&d networks: Theory, empirics, and policy implications. *Review of Economics and Statistics* 101 (3):476–491.

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

Lang, Larry HP and Rene M Stulz. 1994. Tobin's q, corporate diversification, and firm performance. *Journal of Political Economy* 102 (6):1248–1280.

Lerner, Josh and Amit Seru. 2022. The use and misuse of patent data: Issues for finance and beyond. *The Review of Financial Studies* 35 (6):2667–2704.

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

López, Ángel L and Xavier Vives. 2019. Overlapping ownership, r&d spillovers, and antitrust policy. *Journal of Political Economy* 127 (5):2394–2437.

Lucas, Robert E. 1978. On the size distribution of business firms. *The Bell Journal of Economics* :508–523.

Lucking, Brian. 2018. Do R&D tax credits create jobs? Stanford Working Paper.

Lucking, Brian, Nicholas Bloom, and John Van Reenen. 2019. Have r&d spillovers declined in the 21st century? *Fiscal Studies* 40 (4):561–590.

Maksimovic, Vojislav and Gordon Phillips. 2002. Do conglomerate firms allocate resources inefficiently across industries? theory and evidence. *The Journal of Finance* 57 (2):721–767.

Mankiw, N Gregory, David Romer, and David N Weil. 1992. A contribution to the empirics of economic growth. *The Quarterly Journal of Economics* 107 (2):407–437.

Marx, Matt, Jasjit Singh, and Lee Fleming. 2015. Regional disadvantage? employee non-compete agreements and brain drain. *Research Policy* 44 (2):394–404.

McKeown, Greg. 2020. *Essentialism: The disciplined pursuit of less*. Currency.

Milgrom, Paul and John Roberts. 1990. The economics of modern manufacturing: Technology, strategy, and organization. *The American Economic Review* :511–528.

Mullainathan, Sendhil and David Scharfstein. 2001. Do firm boundaries matter? *The American Economic Review* 91 (2):195–199.

Nanda, Ramana and Jesper B Sørensen. 2010. Workplace peers and entrepreneurship. *Management Science* 56 (7):1116–1126.

Phillips, Gordon M and Alexei Zhdanov. 2012. R&D and the incentives from merger and acquisition activity. *The Review of Financial Studies* 26 (1):34–78.

Puri, Manju and Rebecca Zarutskie. 2012. On the life cycle dynamics of venture-capital-and non-venture-capital-financed firms. *The Journal of Finance* 67 (6):2247–2293.

Rhodes-Kropf, Matthew and David T Robinson. 2008. The market for mergers and the boundaries of the firm. *The Journal of Finance* 63 (3):1169–1211.

Robinson, David T. 2008. Strategic alliances and the boundaries of the firm. *Review of Financial Studies* 21 (2):649–681.

Romer, Paul M. 1990. Endogenous technological change. *Journal of Political Economy* 98 (5, Part 2):S71–S102.

Rosen, Sherwin. 1986. The theory of equalizing differences. *Handbook of labor economics* 1:641–692.

Sampat, Bhaven N. 2018. A survey of empirical evidence on patents and innovation. *NBER Working Paper* (w25383).

Schoar, Antoinette. 2002. Effects of corporate diversification on productivity. *The Journal of Finance* 57 (6):2379–2403.

Schumpeter, Joseph A. 1911. *Theorie der wirtschaftlichen entwicklung*. Leipzig: Duncker & Humblot.

Serafinelli, Michel. 2019. Good firms, worker flows, and local productivity. *Journal of Labor Economics* 37 (3):747–792.

Seru, Amit. 2014. Firm boundaries matter: Evidence from conglomerates and R&D activity. *Journal of Financial Economics* 111 (2):381–405.

Sørensen, Jesper B. 2007. Bureaucracy and entrepreneurship: Workplace effects on entrepreneurial entry. *Administrative Science Quarterly* 52 (3):387–412.

Stuart, Toby E and Waverly W Ding. 2006. When do scientists become entrepreneurs? The social structural antecedents of commercial activity in the academic life sciences. *American Journal of Sociology* 112 (1):97–144.

Tate, Geoffrey and Liu Yang. 2015. The bright side of corporate diversification: Evidence from internal labor markets. *The Review of Financial Studies* 28 (8):2203–2249.

Van Reenen, John. 2020. Innovation and human capital policy. In *Innovation and public policy*. University of Chicago Press.

Williamson, Oliver E. 1975. Markets and hierarchies. *New York* 2630.

Zacchia, Paolo. 2019. Knowledge spillovers through networks of scientists. *The Review of Economic Studies* 0:1–30.

Table 1: Summary Statistics

	Mean	Quasi-median	Standard deviation
Firm-year Level Variables			
Made corporate VC investments$_t$ (%)	3.8		
Had $\geq$ 1 patent$_{t-10,t}$ (%)	60.1		
Diversified$_{t-1}$ (%)	78.9		
R&D/Total Assets$_{t-1}$	0.085	0.052	0.102
Log R&D$_{t-1}$	2.53	2.45	2.25
Tobin's Q$_{t-1}$	2.12	1.65	1.59
Firm age$_{t-1}$	20.03	21.03	6.18
Total assets$_{t-1}$ ('000s)	3,483	529	12,630
Firm employment$_{t-1}$	6,107	1,987	12,690
Establishment-year Level Variables			
High-tech industry (%)	64.1		
Employment$_t$	329	122	1,698
Employee entrepreneurship$_{t+3}$ (%)	1.3	0.008	0.024
# employee-founded startups$_{t+3}$	1.15	0.78	1.91
Stayers$_{t+3}$ (%)	47.8	0.523	0.260
Movers to incumbent firms ($\geq$4 yrs old)$_{t+3}$ (%)	26.3	0.225	0.181
Movers to new firms ($\leq$3 yrs old)$_{t+3}$ (%)	3.2	0.020	0.052
Movers to young firms ($<$10 yrs old)$_{t+3}$ (%)	7.2	0.056	0.076
Movers to high-tech young firms ($<$10 yrs old)$_{t+3}$ (%)	4.3	0.031	0.059
Depart employment data$_{t+3}$ (%)	12.4	0.111	0.078
Establishment-founded Startup Level Variables			
Incorporated (%)	69.8		
High-tech industry (%)	49.4		
Exit in 5 years (%)	52.5		
Ever received VC (%)	2.0		
Initial employment	11.83	5.41	29.85
Initial payroll ('000s)	394	119	1,157

NOTE.- This tables shows summary statistics at the firm-year level (top set, 10,500 observations), at the establishment-year level (middle set, 36,000 observations), and at the employee-founded startup level (bottom set, 108,000 observations). We do not show the median or standard deviation for indicators. Since Census disclosure procedures prohibit disclosure of percentile value, we approximate median with a quasi-median, which is estimated using a 99% weight on observations in the interquartile range and a 1% weight on the remaining observations. "Initial" refers to the first year. Payroll is in thousands of dollars.

Table 2: Effect of R&D on Employee Entrepreneurship

Dependent variable: Employee entrepreneurship$_{t+3}$

Panel 1

	(1)	(2)	(3)	(4)	(5)
Firm log R&D$_{t-1}$	0.096**	0.105**	0.106**	0.099*	0.109*
	(0.045)	(0.050)	(0.051)	(0.052)	(0.060)
Baseline Controls		Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	
Firm FE	Yes	Yes	Yes	Yes	Yes
State FE			Yes	Yes	
Industry (SIC3) FE			Yes		
Industry (SIC4) FE				Yes	
Industry (SIC3)-year FE					Yes
State-year FE					Yes
N	36,000	36,000	36,000	36,000	36,000
Adj. R^2	0.156	0.167	0.176	0.184	0.180

Panel 2

	(1)	(2)	(3)	(4)	(5)
Firm log R&D$_{t-1}$	0.102**	0.104**	0.107**	0.104**	0.103**
	(0.052)	(0.051)	(0.051)	(0.051)	(0.057)
Baseline Controls	Yes	Yes	Yes	Yes	Yes
Employee Controls		Yes			
Patent Controls			Yes		
Market R&D Controls				Yes	
Year FE	Yes	Yes	Yes	Yes	Yes
Firm FE	Yes	Yes	Yes	Yes	Yes
State FE		Yes	Yes	Yes	
Industry (SIC3) FE		Yes	Yes		
Industry (SIC4) FE	Yes			Yes	
Industry-state-year FE					Yes
N	36,000	36,000	36,000	36,000	36,000
Adj. R^2	0.181	0.179	0.176	0.176	0.165

NOTE.- This table shows the effect of corporate R&D on employee entrepreneurship (see Section 3.2 for definitions). The sample is an establishment-year panel of public firms. Baseline controls are establishment log employment and log average wage in year t, and firm age, diversified indicator, sales growth, ROA, investment/total assets, Tobin's Q, total assets, PPE investment/total assets, cash/total assets, leverage in year $t-1$. Employee controls are establishment employee average age, female, white, foreign, education, tenure, and experience in year t. Patent activity: firm log patent classes, patents, forward citations, and backward citations in year t. Market R&D Controls are market R&D and market R&D^2 in year $t-1$. Standard errors are clustered by firm. $*$, $**$, and $***$ denote significance at the 10%, 5%, and 1% levels.

Table 3: Parent Firm Variation in Effect of R&D on Employee Entrepreneurship

Dependent variable: Employee entrepreneurship$_{t+3}$

	(1)	(2)	(3)	(4)	(5)
Firm log R&D$_{t-1}$	0.046	0.098*	0.012	0.102**	0.105**
	(0.056)	(0.052)	(0.062)	(0.052)	(0.051)
Firm log R&D$_{t-1}$·Establishment high tech industry	0.083***				
	(0.029)				
Firm log R&D$_{t-1}$·Firm high patent generality$_{t-1}$		0.026*			
		(0.016)			
Firm log R&D$_{t-1}$·Firm large$_{t-1}$			0.133**		
			(0.056)		
Firm log R&D$_{t-1}$·Firm IPO$_{t-3,t-1}$				0.074	
				(0.057)	
Firm log R&D$_{t-1}$·CVC Investment$_t$					-0.011
					(0.046)
Controls	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes
Firm FE	Yes	Yes	Yes	Yes	Yes
State FE	Yes	Yes	Yes	Yes	Yes
Industry (SIC3) FE	Yes	Yes	Yes	Yes	Yes
N	36,000	36,000	36,000	36,000	36,000
Adj. R^2	0.177	0.176	0.176	0.176	0.176

NOTE.- This table shows how the effect of corporate R&D on employee entrepreneurship varies by parent firm characteristics. The sample is an establishment-year panel of public firms. High tech is 1 if the parent establishment is in a high-tech industry, and 0 if it is not. High patent generality is 1 if the parent has above-median patent generality (calculated at the industry-year level), and 0 if it is below median. Large is 1 if the parent firm has above-median total assets (calculated at the firm-year level), and 0 if assets are below median. IPO equals 1 if the firm went public in the past three years, and 0 otherwise. CVC investment is 1 if parent firms that are engaged in corporate venture capital (CVC). The dependent variable is the percent of an establishment's workers as of 1st quarter of year zero who are entrepreneurs as of the first quarter of year three. An entrepreneur is defined as a member of a firm's founding team: a person at a firm no more than 3 years old who is among the top 5 earners at that new firm. All specifications include the indicator variables that are used to interact with R&D (not reported). Controls are establishment log employment and average wage in year t, firm age, diversified indicator, sales growth, ROA, investment/total assets, Tobin's Q, total assets, PPE investment/total assets, cash/total assets, leverage in year $t-1$, and market-level R&D and market-level R&D squared in year $t-1$. Standard errors are clustered by firm. $*$, $**$, and $***$ denote significance at the 10%, 5%, and 1% levels.

Table 4: Instrumented Effect of R&D on Employee Entrepreneurship

Dependent variable:	Firm Log R&D$_{t-1}$		Employee entrepreneurship$_{t+3}$	
	First Stage		Second Stage	
	(1)	(2)	(3)	(4)
Firm log federal R&D tax price$_{t-1}$	-1.960***	-1.464***		
	(0.285)	(0.222)		
Firm log state R&D tax price$_{t-1}$	-1.102*	-0.968**		
	(0.664)	(0.467)		
Instrumented firm log R&D$_{t-1}$			0.534***	0.624**
			(0.203)	(0.270)
Market R&D$_{t-1}$	0.638***	0.416**	0.092	0.165
	(0.229)	(0.165)	(0.437)	(0.432)
Market R&D Squared$_{t-1}$	-0.396**	-0.301**	-0.474	-0.523
	(0.172)	(0.128)	(0.347)	(0.350)
Baseline Controls		Yes		Yes
Year FE	Yes	Yes	Yes	Yes
Firm FE	Yes	Yes	Yes	Yes
State FE	Yes	Yes	Yes	Yes
Industry (SIC3) FE	Yes	Yes	Yes	Yes
N	36,000	36,000	36,000	36,000
F-test (instruments)	24.8	23.0		

NOTE.- This table shows the effect of instrumented R&D on employee entrepreneurship. The sample is an establishment-year panel of public firms. The first stage predicting R&D with the instruments is shown in columns 1 and 2, and the second stage in columns 3 and 4. Employee entrepreneurship is the percent of an establishment's workers as of 1st quarter of year zero who are entrepreneurs as of 1st quarter of year three. An entrepreneur is defined as a member of a firm's founding team: a person at a firm no more than 3 years old who is among the top 5 earners at that new firm. Baseline Controls are establishment log employment and average wage in year t, and firm age, diversified indicator, sales growth, ROA, investment/total assets, Tobin's Q, total assets, PPE investment/total assets, cash/total assets, and leverage in year $t-1$. Standard errors are clustered by firm. $*$, $**$, and $***$ denote significance at the 10%, 5%, and 1% levels.

Table 5: Effect of R&D on Employee Entrepreneurship by Startup Characteristics

Panel 1: Employee-founded startup characteristics

Dependent Variable:	Received VC	Incorporated	High-tech Industry	Log Average Wage_{t+3}	Exit in 5 Years_{t+5}	Same Industry (SIC2) as Parent	Same State as Parent
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Firm log $\text{R\&D}_{t-1}$	0.007***	0.009***	0.010***	0.029***	0.006**	-0.007*	0.002
	(0.001)	(0.003)	(0.004)	(0.007)	(0.003)	(0.003)	(0.002)
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes
State-year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Industry-year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
N	108,000	108,000	108,000	108,000	108,000	108,000	108,000
Adj. R^2	0.079	0.080	0.102	0.318	0.083	0.206	0.053

Panel 2: Input-output relationship between parent firm and startup industries

Dependent variable: Indicator for being in top 5% of closeness distribution

	Parent upstream, employee-founded startup downstream		Parent downstream, employee-founded startup upstream	
Supply chain closeness measure:	Downstream closeness	Upstream closeness	Downstream closeness	Upstream closeness
	(1)	(2)	(3)	(4)
Firm log $\text{R\&D}_{t-1}$	0.008**	-0.003**	0.001	0.006*
	(0.003)	(0.001)	(0.001)	(0.003)
Controls	Yes	Yes	Yes	Yes
State-year FE	Yes	Yes	Yes	Yes
Industry-year FE	Yes	Yes	Yes	Yes
N	108,000	108,000	108,000	108,000
Adj. R^2	0.195	0.115	0.035	0.158

NOTE.- This table studies employee entrepreneurship at the employee-founded startup level. Panel A shows the effect of R&D on startup characteristics. Based on the main variable used in Table 2, we identify whether the new firm associated with the departing employee has a given characteristic: has received VC investments (either before or after the employee-founded startup is identified in year t + 3), is an incorporated business, is in a high-tech industry, exited by year 5 (mostly firm failures), in same two-digit SIC code as the parent establishment, in the same state as the parent establishment. These form binary dependent variables. The exception is column 4, where we consider the departing employee entrepreneur's log wages at the new firm in the 1st quarter of year three. Panel B shows the effect of R&D on employee entrepreneurship based on the supply chain relationship between the parent and the employee-founded startup. The dependent variable is an indicator for the parent-startup pair having a measure of supply chain industry closeness that is in the top 5% of the overall closeness distribution across all parent-startup pairs. In columns 1 and 3, the measure is downstream closeness (downstream industry's share of upstream industry's product). In columns 2 and 4, the measure is upstream closeness (the upstream industry's share of what the downstream industry uses). In columns 1 and 2, the parent is assigned to the upstream industry and the employee-founded startup to the downstream industry (vice versa for columns 3 and 4). In all columns, Controls include the Baseline Controls (included and defined in Table 3), worker-level controls (worker age, worker age squared, female, white, foreign-born, born-in-state, education, years of working experience, years of tenure with the parent employer, log wage at the parent employer before departure), and startup characteristics (age, employment at founding). Column 5 of Panel A also includes control for the startup being VC-backed. Standard errors are clustered by parent firm. $*$, $**$, and $***$ denote significance at the 10%, 5%, and 1% levels.