Entrepreneurial Spillovers from Corporate R&D

Tania Babina and Sabrina T. Howell*

Abstract

This paper offers the first study of how changes in corporate R&D investment affect labor mobility. We document that increases in firm R&D have no measurable effect on employee mobility to other incumbent firms or on exit from employment, but do spur employee departures to join the founding teams of startups. These startups are more likely to be outside the R&D-investing employer's industry, suggesting that the ideas moving via employees to startups would impose diversification costs on the parent. These startups also likely generate substantial spillover benefits, as they are more likely to be VC-backed, high-tech, and high-wage.

*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 and therefore to the productivity and industrial organization literatures (Bernstein & Nadiri 1989, Bloom et al. 2013).[1] 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, particularly outside manufacturing, 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). Employees, whose human capital is inalienable and portable (Aghion & Tirole 1994), are an additional channel for new knowledge and skills to leave the firm.

From a theoretical perspective, it is not obvious how R&D investment will affect labor mobility. 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). Beyond this null hypothesis, we consider four alternatives, not all of which are beneficial for employees. First, retention might increase if R&D generates growth and creates more internal opportunities. Second, R&D investment may lead to automation or other structural changes that make labor redundant or skills obsolete, causing layoffs (Milgrom & Roberts 1990, Acemoglu & Restrepo 2018). Third, R&D investment could create new skills or ideas that improve employees' outside options at other incumbent firms (Herkenhoff et al. 2018). Finally, R&D investment may generate skills or ideas that are specifically valuable for startups (Hart & Moore 1990, Gromb & Scharfstein 2002).

This paper's main contribution is to provide the first direct evidence on the relationship between firm R&D investment and worker mobility. To do this, we construct unique 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

[1]For example, economic growth theories rely on knowledge spillovers (Romer 1990; Eeckhout & Jovanovic 2002; Luttmer 2007; Atkeson & Burstein 2010).

spillovers——not only to other incumbent firms, but also to 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 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 Google between periods of higher and lower R&D, rather than comparing employee departures between Google and Walmart.

Interestingly, we find no effects on departures to other incumbent firms, no effects on exits from the employment sample (which mainly represent unemployment and departures from the labor force, capturing a potential dark-side of R&D), and no effects on employee retention.[3] We find weak positive effects of R&D on employee departures to young firms, which points us in the direction of entrepreneurial outcomes. Indeed, 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.[4] 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 parent firms (e.g. Amazon's headquarters rather than its warehouses). The effect is not driven by recent hires, who might have been hired as a result 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.

[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 & Lerner (2000), Foster, Haltiwanger & Syverson (2008), Gennaioli, La Porta, Lopez-de Silanes & Shleifer (2012), Haltiwanger, Jarmin & Miranda (2013) Decker et al. (2014), and Glaeser, Kerr & Kerr (2015).

[3]The attrition rates in our employment sample are similar to those in the nationally representative Current Population Survey (see Internet Appendix Section A.3).

[4]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 & Kerr (2017), Babina (Forthcoming), and Azoulay, Jones, Kim & Miranda (2018).

While the data do not permit us to pin down a precise channel for R&D-induced entrepreneurial departures, 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 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 & Hart 1986, Aghion & 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.[5] While transaction costs are lower within the boundary of the firm (Atalay, Hortaçsu, Li & Syverson 2019), there is relatively less benefit to locating within the same firms assets that are not complementary (Williamson 1975, Hart & Moore 1990). Consistent with this channel, we document 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.

Firms may permit employees to depart with R&D outputs in order to access the best talent and induce optimal effort. 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. Consistent with this, 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, Celik & Greenwood 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.[6] These ideas benefit from the high-powered incentives that exist in

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

[6]This is due to contracting frictions, as described in Gromb & Scharfstein (2002), Robinson (2008), and Frésard,

small, focused firms (Rhodes-Kropf & Robinson 2008, Phillips & Zhdanov 2012) and are also more conducive to spillovers and growth. Indeed, we document that higher parent R&D is strongly associated with VC backing among employee-founded startups. 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 & Zarutskie 2012). We find that 2% of startups in our sample ever receive VC funding, which is 18 times the national average. Moreover, we 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.[7] 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.

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 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

Hoberg & Phillips (2017).

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

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] The instrumental variables (IV) analysis also finds that R&D has no effect on non-entrepreneurial employee departures or on employee retention, but finds a robust, positive and significant effect on employee-founded startups.

By focusing on innovation inputs and labor reallocation across firms, this paper extends the empirical literature on R&D spillovers, which includes Jones & Williams (1998) and Kerr & Kominers (2015). One line of research focuses on the role of competitive incentives (Atkeson & Burstein 2019, López & Vives 2019), another on spillovers within firms (Herkenhoff et al. 2018, Bilir & Morales 2020), and a large strand on spatial aspects of spillovers (Jaffe, Trajtenberg & Henderson 1993, Griffith, Harrison & Van Reenen 2006, Kantor & Whalley 2014). In a recent important empirical contribution, Bloom et al. (2013) show that R&D spillovers dominate the negative competition effect that R&D can impose through business stealing. They also suggest it is important to study the channels through

[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 & Gilson (1999), Curtis & Decker (2018), Lucking (2018)

which R&D spillovers operate, including labor mobility. 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. Surprisingly, we find no effects of R&D on departures to other incumbent firms, but find a strong relationship between R&D and reallocation to startups, with cross-sectional evidence consistent with R&D-induced startups being a source of spillovers. Acemoglu et al. (2013) argue that R&D subsidies may be misguided because they favor incumbents at the expense of entrants. If entrepreneurial spillovers from corporate R&D were included in their model, the policy implications might be different.

Our finding helps explain why high-growth startup founders are often former employees of large incumbent firms (Gompers, Lerner & Scharfstein 2005, Babina Forthcoming). We offer corporate R&D as a new source for where ideas and human capital for the entrepreneurial sector come from (Aghion & Jaravel 2015, Akcigit & Ates 2019). There is a rich strategy literature studying the origins of employee entrepreneurship (Klepper 2009). However, Agarwal & Shah (2014) highlight important limitations with these analyses; they are based on cross-sectional, single-industry, or case study data and generally have no causal interpretation. Our data and empirical design overcome these limitations and ask a more general question of how R&D investment affects labor reallocation. We find that an increase in firm R&D investment predicts departures to startups in *different* industries, which are missing in existing literature that focuses on within-industry spinoffs (Agarwal & Shah 2014).

Other related work on knowledge diffusion has emphasized the importance of inventor networks and collaboration among scientists (Borjas & 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. For example, an employee of a firm creating maps may use technology underlying the parent's mapping software but apply it to agriculture. Employee entrepreneurship to other sectors is a channel for knowledge spillovers and technological interconnections across sectors.

2 Hypothesis development

This section presents four not mutually exclusive or exhaustive hypotheses about how innovation investment interacts with labor mobility. Specifically, corporate R&D may have a positive effect on employee (1) Retention; (2) Departures from the labor force; (3) Mobility to other firms; and (4) Entrepreneurship. As we are focused on how R&D affects exposed human capital, we do not consider new employees. This section concentrates on (4), as it has the richest theoretical motivation and is the only one with empirical support. We further offer two predictions about (a) the type of R&D-generated growth option that is most likely to locate in a startup; and (b) the circumstances under which an idea may be more valuable in a standalone firm founded by the employee than in the parent firm.

(1) Corporate R&D and employee retention. The purpose of R&D is to create growth options for the firm and, potentially, to steal business from competitors (Bloom et al. 2013). If R&D leads to growth, it may create better job advancement opportunities. It could also make the firm a more interesting place to work; that is, it might increase workplace amenities (Rosen 1986). For example, R&D may lead to new product development that is stimulating for the employee. In this case, R&D may make existing workers less likely to leave, leading to a positive effect on employee retention. If the data support this mechanism, it would stand in striking contrast to the common, albeit largely untested, perception that R&D increases knowledge diffusion through labor mobility.

(2) Corporate R&D and departures from the labor force. R&D may lead to automation or other structural changes that make labor redundant or skills obsolete, causing layoffs (Brynjolfsson & McAfee 2014). Even if innovation investment is associated with new labor-intensive tasks that are complementary with labor-replacing technology, as in Acemoglu & Restrepo (2018), we may expect that some individuals will not have the requisite skills to transition to the new tasks, and automation or digitization will lead to an increase in departures to unemployment. Finding a positive effect of R&D on departures from the labor force would shed light on the decline in the low-skill labor share of national income. However, it may be that automation is reflected in capital expenditure, and does not have a direct relationship to R&D.

An alternative mechanism for R&D to increase layoffs is if it is associated with restructuring. When firms invest in new technologies, they increase R&D and tend to restructure their entire workforce—the supermodularity concept of Milgrom & Roberts (1990). R&D outputs could cause the firm to pursue a new strategic direction and find some existing workers no longer useful. For example, GM's decision to lay off more than 14,000 employees in North America in late 2018 was in part attributed to its new focus on electric vehicles and autonomous driving technology.[10]

(3) Corporate R&D and employee mobility to incumbent firms. It is well known that knowledge diffuses between firms (Griliches 1992). The R&D process is human capital-intensive, requiring creative effort on the part of employees. This innovation effort is hard or impossible to contract on ex-ante, and hard or impossible to verify ex-post (Grossman & Hart 1986 and Aghion & Tirole 1994). These contracting frictions could lead R&D to induce more employee mobility to other incumbent firms. That is, for some employees R&D may create better outside options at peer firms. This could involve 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.[11]

(4) Corporate R&D and employee departures to entrepreneurship. 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 & 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, Hoberg & Phillips (2017) model these frictions and conclude that control rights should be allocated to

[10] See here.

[11] For example, Singh & Agrawal (2011) find that when inventors move between firms, they make use of their existing inventions, which benefits the hiring firm. Similarly, Tambe & Hitt (2013) find that productivity spillovers occur when IT workers move between firms. Other literature on labor mobility and knowledge spillovers includes Almeida & Kogut (1999), Agrawal, Cockburn & McHale (2006), Breschi & Lissoni (2009), and Moretti & Wilson (2017). Finally, individuals are also known to play an important role in international knowledge flows (Oettl & Agrawal 2008, Docquier & Rapoport 2012, and Poole 2013).

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 as a means 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, Celik & Greenwood (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. (2019) 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.

(4a): R&D-induced employee-founded startups are 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 & 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.

(4b): 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.[12] 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 & Moore 1990). Empirical work finds a negative correlation between firm performance and diversification (Lang & 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 2012). 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 venture capital, 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. Both hypotheses 4(a) and 4(b) are 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 each source of data and explains the key variables we use in our analysis. It also discusses potential concerns with the data.

[12] As one example, consider the agricultural insurance startup startup WeatherBill (later renamed The Climate Corporation and ultimately acquired by Monsanto). Its founders, David Friedberg and Siraj Khaliq, left Google in 2006 to launch the new firm, which used artificial intelligence insights from Google to better price insurance. See here. Google may be more permissive of employee entrepreneurship in areas such as insurance, which are relatively far from its main focus. In contrast, Google filed a lawsuit when one of its key autonomous vehicle employees departed to Uber's autonomous vehicle unit.

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.[13] 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 longitudinally links establishments that are part of the same firm. Incorporated businesses rather than sole proprietorships or partnerships comprise about 83 percent of the LBD.[14] For further details about the LBD, see Jarmin & 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 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 & 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 Internet Appendix Section A.3, and our key independent variable—R&D—does not predict departures out of sample (Table 8 column 6), suggesting that

[13] 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 & Seru 2017). 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.

[14] 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.

incomplete coverage likely does not affect our results.[15] 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.[16] 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.[17] 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.

More systematically, 60 percent of U.S. employment, or about 150 million people, and 52 percent of U.S. inventors are located in our states.[18] 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. Our data are also representative at the industry level, as shown in Appendix Table A.2.[19] For example, in the 2002-08 period, Professional and Business Services (which includes information technology

[15]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).

[16]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.

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

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

[19]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.

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." We do the linkage in the first quarter of each year since the annual LBD measures employment and payroll in March. We drop establishments 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 & 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 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 Internet Appendix Table A.1.

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

[20]We obtain similar results if we drop establishments with less than five or less than 15 employees.

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.

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 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. (2018) find that a firm's top three initial earners usually include the firm's owners. This is because the W-2 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.[21] 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 (Forthcoming).[22] 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

[21] See here.

[22] 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.

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 Panels 1-3 contain 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.[23] Our dependent variables are measured at the establishment-year level as explained above (Panel 2). 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, Kerr & Nanda (2015) find in LBD/LEHD matched data that 1.7 percent of workers become entrepreneurs over a four-year period. Panel 3 of Table 1 describes 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 & 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

[23] 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.

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.

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:

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

Our primary independent variable is log R&D among firms with positive R&D (the results are robust to including non-R&D performing firms and using R&D/Assets). We employ firm fixed effects to control for time-invariant differences across firms. This means that 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 Google with those from Walmart, within-firm estimation allows us to compare employee departures within Google between periods of higher and lower R&D. Industry-year fixed effects at the establishment level 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 mis-measurement. 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. We control for 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. 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 assets), size (log total assets), cash holdings, age, and diversification (an indicator for firms having establishments in multiple SIC three-digit industries).

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. The data do not actually support this concern, but we explore whether it may affect the results using the best available instruments for R&D investment, which are provided by Bloom et al. (2013) and are based on changes in the tax price of R&D (also used in König et al. (2019)).

Specifically, one instrument 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:

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

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.4.

5 Results

In Section 5.1, we test *Hypotheses 1-3*. We present the main results for *Hypothesis 4* – the effect of R&D on employee entrepreneurship – in Sections 5.2 and 5.3. Alternative measures of entrepreneurship and R&D, as well as heterogeneity in parent firm characteristics, are considered in Section 5.4 and 5.5. Further robustness tests are in Section 5.6.

5.1 Effects on retention, layoffs, and mobility to incumbents

We begin by testing whether R&D affects retention, departures from the labor force, or mobility to other incumbent firms. The effects of R&D on these outcomes are reported in Table 2 using the OLS and IV approaches. We focus on the OLS results, as this is our preferred model for entrepreneurship (discussed below), though the results are not qualitatively different using the IV. Columns 1 and 2 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 1 of -1.13 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. Columns 3 and 4 show no relationship between R&D and employee exits from employment; we can rule out effects outside the 95 percent confidence interval bounds of -2.2 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. Columns 5 and 6 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, suggesting that neither innovation inputs (R&D) nor outputs (patents) are related to non-entrepreneurial labor reallocation.

We do find positive effects on moves to young firms. Column 7 of Table 2 indicates that a 100 percent increase (about one standard deviation) in parent firm R&D is associated with an 4.8 percent increase in the employee departure rate to incumbent firms less than 10 years old, relative to the sample mean. The IV estimate is larger, but less precise (column 8). When we require the young incumbent firm to be high tech, the coefficient of 0.23 represents a 5.4 percent increase relative to the sample mean (column 9). The IV result has no significance (column 10). Together, these results show that R&D has no measurable effect on labor reallocation except to firms that are close to startups, pointing us toward entrepreneurial departures.

5.2 Effect on employee entrepreneurship

Table 3 presents the relationship between parent firm R&D and employee departures to entrepreneurship. In Column 1, we start with the simplest specification with firm and year fixed effects. Higher R&D predicts a statistically significant increase in entrepreneurial departures. We next examine whether technology or other investment opportunity shocks lead to both R&D and entrepreneurial departures—this is the main concern in our setting. In Column 2, we add controls that may be correlated with such shocks, such as establishment employment and payroll as well as firm sales growth, profitability, investments, Tobin's Q, and cash. We also control for variables that can change through a firm's lifecycle and could be correlated with R&D, such as firm age, diversification, total assets, fixed assets, and leverage. However, the coefficient on R&D does not attenuate with these additional controls, and, in fact, slightly increases. This holds true when we further control for industry and state fixed effects in Columns 3 and 4 (Column 3 replicates the specification reported in Table 2). While most controls have no predictive power for entrepreneurship, larger establishments have less employee entrepreneurship, consistent with prior work (Nanda & Sørensen 2010). Finally, in Column 5 we add industry-year and state-year fixed effects, which should be correlated with technology shocks within industries and locations. Again, the coefficient does not attenuate.[24]

The relationship between R&D and employee entrepreneurship is economically meaningful. From Column 5, 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

[24]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 2.

rate of employee departures to entrepreneurship, relative to the sample mean of 1.3 percent.[25] The result is robust to a wide array of alternative controls and fixed effects, shown across the eight models in Panels 1 and 2.[26] Column 2 of Panel 2 employs additional establishment-level controls. Establishments with a higher share of white workers or foreign-born workers are associated with more employee entrepreneurship. The results do not attenuate with wage controls, so the effect is not driven by an increase in employee wages.

We consider alternative measures of R&D in Table 4. When the independent variable is an indicator for an above median change in R&D, the effect is .089, 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 is 7.7 percent of all employee-founded startups in the data.[27] Table 4 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.[28]

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. Table 1 Panel 4A shows that individuals moving to

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

[26]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 number of coefficients we may disclose. The controls are at the firm level, except for employment and payroll 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.

[27]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.

[28]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.

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. Table 1 Panel 4B shows that the parent (R&D-performing) employing firm has similar characteristics across the two groups, and Panel 4C 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.

5.3 Instrumented effect on employee entrepreneurship

While an OLS strategy does not permit us to completely rule out omitted variable bias, the evidence in Section 5.2 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 5 contains the instrumented results using the same main specifications from Table 3. The coefficients are statistically significant and larger than the OLS results.[29] Our preferred specification, in column 5, 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. 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 & 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

[29] 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 3 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.

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 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 Alternative employee entrepreneurship measures

To demonstrate that our entrepreneurship result is not sensitive to a particular construction of the outcome variable, we consider alternative measures of employee entrepreneurship in Table 6. 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 $t = 0$. 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 (Table 6 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 (Table 6 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 3-4, we consider only startups founded by year $t + 2$. We continue to find a positive, significant coefficient using this more immediate measure. We then consider one-year-old startups in year $t + 2$. 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.

Last, 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.5 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.

5.5 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 located 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.[30] 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 7 column 1).[31] 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. This result also suggests that if the effect is different in California and Massachusetts, it is likely higher.

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 7 column 2). Also firms that patent in more classes tend to have higher employee entrepreneurship rates (Table 3 Panel 2 column 5). Thus, consistent with the spillover interpretation, firms doing broader research have more employee-founded startups per dollar of R&D.

5.6 Additional robustness tests

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.6. 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.

Another possible source of endogeneity is that R&D may lead the firm to hire new employees,

[30] 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.

[31] We do not use the IV estimator in this table because there is insufficient power to identify the interaction term of interest.

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.[32] Therefore, hiring related to the increase in R&D does not drive the effect.

6 Mechanisms

This section considers evidence for *Hypotheses 4a* (Section 6.1) and *4b* (Section 6.2), as stipulated in Section 2. In Section 6.3, we discuss how our patent results reflect incomplete contracting, which is crucial to our proposed mechanism. Evidence against alternative mechanisms is in Section 6.4, though we do not claim that these are entirely non-operative.

6.1 High-risk high-growth

Our first test of *Hypothesis 4a* concerns VC backing. VC-backed startups are widely known to be risky, associated with new-to-the-world ideas, and potentially high growth (Kaplan & Lerner 2010, Gornall & Strebulaev 2015)—the type of startups that are an important source of spillovers and growth. In Table 8 we examine whether parent R&D is associated with certain startup characteristics.[33] Here, analysis is conducted at the startup level and the sample consists of all employee-founded startups in our data. The dependent variable in Table 8 Panel 1 column 1 is one if the employee-founded startup receives VC. The independent variables are specific either to the employee, parent firm, or parent firm establishment. 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

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

[33] 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.

is VC-backed. These results provide the first empirical evidence that, among parent firm observables, R&D is by far the strongest predictor of VC-backed, employee-founded startups.

Levine & 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." R&D-induced startups are more likely to be incorporated (Table 8 Panel 2 column 1). 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 (Table 8 Panel 2 column 2). 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 (Table 8 Panel 2 column 3). 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 4, 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 impact, high tech, and high risk, consistent with *Hypothesis 4a*.

6.2 Costly diversification

Hypothesis 4b of a 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 5 of Table 8 Panel 2, 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. [34]

It may initially seem counter-intuitive that R&D leads employees to found firms in different industries. However, consider three 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 (Glaeser 2011). 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 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.[35] 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). Finally, consider the example of WeatherBill from Section 2. Founded by former Google employees, WeatherBill would be in SIC 63 (Insurance Carriers). Google's parent company Alphabet is in SIC 73 (Business Services). In all three 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. These results are consistent with *Hypothesis 4b*.

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

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

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

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.[36] For both measures, a higher value means they are closer.

The results are in Table 8 Panel 3. We first assign the parent to the upstream industry, and 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.[37] 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

[36]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.

[37]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.

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 & 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 citations have an effect on employee entrepreneurship (Table 3 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. The first measure is inflows: for patent classes A and B, this is B's cites of A as a share of the total cites to A. The second measure is outflows: A's cites of B as a share of all the citations from A. We create this measure at the class-year level, and then assign patent classes to industries using the patent-to-SIC concordance developed by Kerr (2008).[38] When we interact R&D with these measures of knowledge sharing, we find no effects. This supports the conclusion that our results reflect R&D output that is not patented. Ellison, Glaeser & Kerr (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

While cross-sectional evidence supports *Hypothesis 4*, it does not rule out alternative mechanisms. This section considers whether the evidence supports four additional channels.

[38]We are especially grateful to Bill Kerr for his help with this exercise.

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 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, Table 3 Panel 1 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 & Ding 2006, Sørensen 2007). Instead, large firms drive the effect (Table 7 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, Ouimet & Zarutskie 2019). In Table 7 column 4 we interact R&D with an indicator for having had an IPO within the last three years. The interaction is positive and insignificant, but 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 & Schankerman 2013) and within-state inventor mobility (Marx, Singh & Fleming 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, a 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.[39] 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

[39]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.

no effect of an interaction between R&D and the parent having a corporate VC program.

7 Conclusion

Human capital is central to modern theories of economic growth (Romer 1990, Jones 2014). Though the literature has focused on schooling (Card 2001, Cunha & Heckman 2007), firms also play a role in human capital formation. Corporate research effort yields ideas that are first embodied in people and then ultimately in new types of capital inputs (Mankiw, Romer & Weil 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. In contrast with common assumptions in the innovation literature, we find no relationship between R&D and labor mobility to other incumbents.

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. Instead, we document a likely unintended consequence of innovation inputs. 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. Consistent with influential theories of the firm, 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.

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 & 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.

References

Abowd, J. M., Stephens, B. E., Vilhuber, L., Andersson, F., McKinney, K. L., Roemer, M. & Woodcock, S. (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, pp. 149–230.

Acemoglu, D., Akcigit, U., Bloom, N. & Kerr, W. (2013), 'Innovation, reallocation and growth', *NBER Working Paper No. 18993* .

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

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

Aghion, P. & Jaravel, X. (2015), 'Knowledge spillovers, innovation and growth', *The Economic Journal* **125**(583), 533–573.

Aghion, P. & Tirole, J. (1994), 'The management of innovation', *The Quarterly Journal of Economics* **109**(4), 1185–1209.

Agrawal, A., Cockburn, I. & McHale, J. (2006), 'Gone but not forgotten: Knowledge flows, labor mobility, and enduring social relationships', *Journal of Economic Geography* **6**(5), 571–591.

Agrawal, A., Rosell, C. & Simcoe, T. S. (2014), Do tax credits affect R&D expenditures by small firms? Evidence from Canada, Technical report, National Bureau of Economic Research.

Akcigit, U. & Ates, S. T. (2019), Ten facts on declining business dynamism and lessons from endogenous growth theory, Technical report, National Bureau of Economic Research.

Akcigit, U., Celik, M. A. & Greenwood, J. (2016), 'Buy, keep, or sell: Economic growth and the market for ideas', *Econometrica* **84**(3), 943–984.

Almeida, P. & Kogut, B. (1999), 'Localization of knowledge and the mobility of engineers in regional networks', *Management science* **45**(7), 905–917.

Angrist, J. D. & Imbens, G. W. (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, E., Hortaçsu, A., Li, M. J. & Syverson, C. (2019), 'How wide is the firm border?', *The Quarterly Journal of Economics* **134**(4), 1845–1882.

Atkeson, A. & Burstein, A. (2019), 'Aggregate implications of innovation policy', *Journal of Political Economy* **127**(6), 2625–2683.

Atkeson, A. & Burstein, A. T. (2010), 'Innovation, firm dynamics, and international trade', *Journal of political economy* **118**(3), 433–484.

Azoulay, P., Jones, B., Kim, J. D. & Miranda, J. (2018), Age and high-growth entrepreneurship, Technical report, National Bureau of Economic Research.

Babina, T. (Forthcoming), 'Destructive creation at work: How financial distress spurs entrepreneurship', *Review of Financial Studies* .

Babina, T., Ouimet, P. P. & Zarutskie, R. (2019), 'Ipos, human capital, and labor reallocation'. Working paper available at https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2692845.

Balsmeier, B., Kurakina, M. & Fleming, L. (2018), 'R&D tax credits: Mechanisms of private and public value'. Working Paper.

Bankman, J. & Gilson, R. J. (1999), 'Why start-ups?', *Stanford Law Review* **51**, 289, 293–95.

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

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

Bernstein, J. I. & Nadiri, M. I. (1989), 'Research and development and intra-industry spillovers: an empirical application of dynamic duality', *The Review of Economic Studies* **56**(2), 249–267.

Bertrand, M., Duflo, E. & Mullainathan, S. (2004), 'How much should we trust differences-in-differences estimates?', *The Quarterly journal of economics* **119**(1), 249–275.

Bilir, L. K. & Morales, E. (2020), 'Innovation in the global firm', *Journal of Political Economy* **128**(4), 000–000.

Billings, A., Glazunov, S. & Houston, M. (2001), 'The role of taxes in corporate research and development spending', *R&D Management* **31**(4), 465–477.

Bloom, N., Griffith, R. & Van Reenen, J. (2002), 'Do R&D tax credits work? Evidence from a panel of countries 1979–1997', *Journal of Public Economics* **85**(1), 1–31.

Bloom, N., Schankerman, M. & Van Reenen, J. (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, G. J. & Doran, K. B. (2015), 'Which peers matter? the relative impacts of collaborators, colleagues, and competitors', *Review of economics and statistics* **97**(5), 1104–1117.

Breschi, S. & Lissoni, F. (2009), 'Mobility of skilled workers and co-invention networks: an anatomy of localized knowledge flows', *Journal of Economic Geography* p. lbp008.

Brown, J. R. & Petersen, B. C. (2011), 'Cash holdings and R&D smoothing', *Journal of Corporate Finance* **17**(3), 694–709.

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

Cappelen, Å., Raknerud, A. & Rybalka, M. (2012), 'The effects of R&D tax credits on patenting and innovations', *Research Policy* **41**(2), 334–345.

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

Chirinko, R. S. & Wilson, D. J. (2008), 'State investment tax incentives: A zero-sum game?', *Journal of Public Economics* **92**(12), 2362–2384.

Chirinko, R. S. & Wilson, D. J. (2011), 'Tax competition among US states: Racing to the bottom or riding on a seesaw?', *Working Paper* .

Clausen, T. H. (2009), 'Do subsidies have positive impacts on R&D and innovation activities at the firm level?', *Structural Change and Economic Dynamics* **20**(4), 239–253.

Collins, J. C. (2009), *How the mighty fall: And why some companies never give in*, Random House.

Cunha, F. & Heckman, J. (2007), 'The technology of skill formation', *American Economic Review* **97**(2), 31–47.

Curtis, E. M. & Decker, R. (2018), Entrepreneurship and state taxation, Technical report, Board of Governors of the Federal Reserve System (US).

Czarnitzki, D., Hanel, P. & Rosa, J. M. (2011), 'Evaluating the impact of R&D tax credits on innovation: A microeconometric study on canadian firms', *Research Policy* **40**(2), 217–229.

Dechezleprêtre, A., Einiö, E., Martin, R., Nguyen, K.-T. & Van Reenen, J. (2016), Do tax incentives for research increase firm innovation? An RD design for R&D, Technical report, National Bureau of Economic Research.

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

Docquier, F. & Rapoport, H. (2012), 'Globalization, brain drain, and development', *Journal of Economic Literature* **50**(3), 681–730.

Eeckhout, J. & Jovanovic, B. (2002), 'Knowledge spillovers and inequality', *American Economic Review* **92**(5), 1290–1307.

Ellison, G., Glaeser, E. L. & Kerr, W. R. (2010), 'What causes industry agglomeration? Evidence from coagglomeration patterns', *The American Economic Review* **100**(3), 1195–1213.

Foster, L., Haltiwanger, J. & Syverson, C. (2008), 'Reallocation, firm turnover, and efficiency: Selection on productivity or profitability?', *American Economic Review* **98**(1), 394–425.

Frésard, L., Hoberg, G. & Phillips, G. (2017), 'Vertical acquisitions, integration and the boundaries of the firm'.

Garcia-Macia, D., Hsieh, C.-T. & Klenow, P. J. (2019), 'How destructive is innovation?', *Econometrica* **87**(5), 1507–1541.

Garmaise, M. J. (2011), 'Ties that truly bind: Noncompetition agreements, executive compensation, and firm investment', *The Journal of Law, Economics, and Organization* **27**(2), 376–425.

Gennaioli, N., La Porta, R., Lopez-de Silanes, F. & Shleifer, A. (2012), 'Human capital and regional development', *The Quarterly Journal of Economics* **128**(1), 105–164.

Glaeser, E. (2011), *Triumph of the city: How urban spaces make us human*, Penguin Press. p. 47.

Glaeser, E. L., Kerr, S. P. & Kerr, W. R. (2015), 'Entrepreneurship and urban growth: An empirical assessment with historical mines', *Review of Economics and Statistics* **97**(2), 498–520.

Goldin, C., Kerr, S. P., Olivetti, C., Barth, E. et al. (2017), 'The expanding gender earnings gap: evidence from the lehd-2000 census', *American Economic Review* **107**(5), 110–14.

Gompers, P., Lerner, J. & Scharfstein, D. (2005), 'Entrepreneurial spawning: Public corporations and the genesis of new ventures, 1986 to 1999', *The Journal of Finance* **60**(2), 577–614.

Gornall, W. & Strebulaev, I. A. (2015), 'The economic impact of venture capital: Evidence from public companies', *Working Paper* .

Griffith, R., Harrison, R. & Van Reenen, J. (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, Z. (1992), 'The search for R&D spillovers', *Scandinavian Journal of Economics* **94**(3).

Gromb, D. & Scharfstein, D. (2002), Entrepreneurship in equilibrium, Technical report, National bureau of economic research.

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

Guceri, I. & Liu, L. (2017), 'Effectiveness of fiscal incentives for R&D: Quasi-experimental evidence'. IMF Working Paper.

Guzman, J. (2019), 'Go west young firm: Agglomeration and embeddedness in startup migrations to silicon valley'.

Hall, B. H. (1993), 'R&D tax policy during the 1980s: Success or failure?', *Tax policy and the economy* **7**, 1–35.

Hall, B. & Van Reenen, J. (2000), 'How effective are fiscal incentives for R&D? A review of the evidence', *Research Policy* **29**(4), 449–469.

Haltiwanger, J., Jarmin, R. S. & Miranda, J. (2013), 'Who creates jobs? small versus large versus young', *Review of Economics and Statistics* **95**(2), 347–361.

Hart, O. & Moore, J. (1990), 'Property rights and the nature of the firm', *Journal of Political Economy* **98**, 1152–53.

Heggedal, T.-R., Moen, E. R. & Preugschat, E. (2017), 'Productivity spillovers through labor mobility in search equilibrium', *Journal of Economic Theory* **169**, 551–602.

Herkenhoff, K., Lise, J., Menzio, G. & Phillips, G. (2018), 'Worker mobility and the diffusion of knowledge'.

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

Jaffe, A. B., Trajtenberg, M. & Henderson, R. (1993), 'Geographic localization of knowledge spillovers as evidenced by patent citations', *the Quarterly journal of Economics* pp. 577–598.

Jarmin, R. S. & Miranda, J. (2002), 'The longitudinal business database'.

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

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

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

Jones, C. I. & Williams, J. C. (1998), 'Measuring the social return to R&D', *Quarterly Journal of Economics* pp. 1119–1135.

Kahn, R. & Whited, T. M. (2017), 'Identification is not causality, and vice versa', *Review of Corporate Finance Studies* **7**(1), 1–21.

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

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

Kaplan, S. N. & Lerner, J. (2016), Venture capital data: Opportunities and challenges, *in* 'Measuring Entrepreneurial Businesses: Current Knowledge and Challenges', University of Chicago Press.

Kerr, S., Kerr, W. R. & Nanda, R. (2015), House money and entrepreneurship, Technical report, National Bureau of Economic Research.

Kerr, S. P. & Kerr, W. R. (2017), 'Immigrant entrepreneurship', *Measuring Entrepreneurial Businesses: Current Knowledge and Challenges* p. 187.

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

Kerr, W. R. & Kominers, S. D. (2015), ‘Agglomerative forces and cluster shapes’, *Review of Economics and Statistics* **97**(4), 877–899.

Kim, J. & Marschke, G. (2005), ‘Labor mobility of scientists, technological diffusion, and the firm’s patenting decision’, *RAND Journal of Economics* pp. 298–317.

Klepper, S. (2009), ‘Spinoffs: A review and synthesis’, *European Management Review* **6**(3), 159–171.

Kondo, J. E., Li, D. & Papanikolaou, D. (2019), ‘Trust, collaboration, and economic growth’.

König, M. D., Liu, X. & Zenou, Y. (2019), ‘R&d networks: Theory, empirics, and policy implications’, *Review of Economics and Statistics* **101**(3), 476–491.

Kortum, S. & Lerner, J. (2000), ‘Assessing the contribution of venture capital to innovation’, *The Rand Journal of Economics* **31**(4), 674.

Lang, L. H. & Stulz, R. M. (1994), ‘Tobin’s q, corporate diversification, and firm performance’, *Journal of Political Economy* **102**(6), 1248–1280.

Lerner, J. & Seru, A. (2017), The use and misuse of patent data: Issues for corporate finance and beyond, Technical report, National Bureau of Economic Research.

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

López, Á. L. & Vives, X. (2019), ‘Overlapping ownership, r&d spillovers, and antitrust policy’, *Journal of Political Economy* **127**(5), 2394–2437.

Lucas, R. E. (1978), ‘On the size distribution of business firms’, *The Bell Journal of Economics* pp. 508–523.

Lucking, B. (2018), ‘Do R&D tax credits create jobs?’. Stanford Working Paper.

Luttmer, E. G. (2007), ‘Selection, growth, and the size distribution of firms’, *The Quarterly Journal of Economics* **122**(3), 1103–1144.

Ma, S. (2016), ‘The life cycle of corporate venture capital’. Working Paper.

Maats, F., Metrick, A., Yasuda, A., Hinkes, B. & Vershovski, S. (2011), ‘On the consistency and reliability of venture capital databases’, *Unpublished working paper* .

Maksimovic, V. & Phillips, G. (2002), ‘Do conglomerate firms allocate resources inefficiently across industries? theory and evidence’, *The Journal of Finance* **57**(2), 721–767.

Mamuneas, T. P. & Nadiri, M. I. (1996), ‘Public R&D policies and cost behavior of the US manufacturing industries’, *Journal of Public Economics* **63**(1), 57–81.

Mankiw, N. G., Romer, D. & Weil, D. N. (1992), ‘A contribution to the empirics of economic growth’, *The Quarterly Journal of Economics* **107**(2), 407–437.

Marx, M., Singh, J. & Fleming, L. (2015), ‘Regional disadvantage? employee non-compete agreements and brain drain’, *Research Policy* **44**(2), 394–404.

McCutchen, W. W. (1993), ‘Estimating the impact of the R&D tax credit on strategic groups in the pharmaceutical industry’, *Research Policy* **22**(4), 337–351.

McKeown, G. (2012), ‘The disciplined pursuit of less’, *Harvard Business Review* .

Milgrom, P. & Roberts, J. (1990), ‘The economics of modern manufacturing: Technology, strategy, and organization’, *The American Economic Review* pp. 511–528.

Miller, C. R. & Richard, B. (2010), ‘The policy diffusion of the state R&D investment tax credit’, *State and Local Government Review* **42**(1), 22–35.

Moretti, E. & Wilson, D. J. (2017), ‘The effect of state taxes on the geographical location of top earners: Evidence from star scientists’, *American Economic Review* **107**(7), 1858–1903.

Mullainathan, S. & Scharfstein, D. (2001), ‘Do firm boundaries matter?’, *The American Economic Review* **91**(2), 195–199.

Nanda, R. & Sørensen, J. B. (2010), ‘Workplace peers and entrepreneurship’, *Management Science* **56**(7), 1116–1126.

Oettl, A. & Agrawal, A. (2008), ‘International labor mobility and knowledge flow externalities’, *Journal of International Business Studies* **39**(8), 1242–1260.

Paff, L. A. (2005), ‘State-level R&D tax credits: A firm-level analysis’, *The BE Journal of Economic Analysis & Policy* **5**(1).

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

Poole, J. P. (2013), ‘Knowledge transfers from multinational to domestic firms: Evidence from worker mobility’, *Review of Economics and Statistics* **95**(2), 393–406.

Puri, M. & Zarutskie, R. (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, M. & Robinson, D. T. (2008), ‘The market for mergers and the boundaries of the firm’, *The Journal of Finance* **63**(3), 1169–1211.

Robinson, D. T. (2008), ‘Strategic alliances and the boundaries of the firm’, *Review of Financial Studies* **21**(2), 649–681.

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

Rosen, S. (1986), ‘The theory of equalizing differences’, *Handbook of labor economics* **1**, 641–692.

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

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

Schumpeter, J. A. (1911), *Theorie Der Wirtschaftlichen Entwicklung*, Duncker & Humblot, Leipzig.

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

Singh, J. & Agrawal, A. (2011), ‘Recruiting for ideas: How firms exploit the prior inventions of new hires’, *Management Science* **57**(1), 129–150.

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

Stuart, T. E. & Ding, W. W. (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.

Tambe, P. & Hitt, L. M. (2013), ‘Job hopping, information technology spillovers, and productivity growth’, *Management science* **60**(2), 338–355.

Tate, G. & Yang, L. (2015), ‘The bright side of corporate diversification: Evidence from internal labor markets’, *The Review of Financial Studies* **28**(8), 2203–2249.

Williamson, O. E. (1975), ‘Markets and hierarchies’, *New York* **2630**.

Wilson, D. J. (2009), ‘Beggar thy neighbor? The in-state, out-of-state, and aggregate effects of R&D tax credits’, *The Review of Economics and Statistic* .

Wilson, D. J. et al. (2005), ‘The rise and spread of state R&D tax credits’, *FRBSF Economic Letter* .

Wu, Y. (2008), ‘State R&D tax credits and high-technology establishments’, *Economic Development Quarterly* **22**(2), 136–148.

Zacchia, P. (2019), ‘Knowledge spillovers through networks of scientists’, *The Review of Economic Studies* **0**, 1–30.

Table 1: Summary Statistics

Panel 1: Firm-year level variables

	Mean	Quasi-median	Standard deviation
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

Panel 2: Establishment-year level variables

	Mean	Quasi-median	Standard deviation
Weak non-compete enforcement (state) (%)	61.3		
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

Note: Panel 1 shows summary statistics at the firm-year level (10,500 observations), and Panel 2 at the establishment-year level (36,000 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.

Panel 3: Employee-founded startup level variables

	Mean	Quasi-median	Standard deviation
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

Panel 4: Comparison between employees who depart to entrepreneurship and those who depart to incumbents

	Workers who join founding teams of startups ($\leq$3 yrs old)	Workers who move to incumbents ($\geq$4 yrs old)
	A. Employee characteristics	
Share female$_t$	33	41
Share white$_t$	80	70
Share foreign$_t$	7.7	7.0
Average education$_t$	13.7	13.5
Average age$_t$	35.2	34.8
Average tenure (at public firm; years)$_t$	2.07	1.99
Average wages (at public firm; '000s)$_t$	57.8	39.7
	B. Public firm characteristics	
Log R&D$_{t-1}$	3.80	3.71
ROA$_{t-1}$	0.17	0.17
Tobin's Q$_{t-1}$	2.19	2.08
Firm age$_{t-1}$	21.7	22.3
Total assets$_{t-1}$ ('000s)	11.2	11.4
	C: Future employer characteristics	
Firm age$_{t+3}$	1.59	20.9
Firm employment$_{t+3}$	14.6	1,340
Firm payroll ('000s)$_{t+3}$	511	44,520

Note: Panel 3 shows summary statistics at the employee-founded startup level. All variables have 108,000 observations. Variables through "Ever received VC" are indicators, and the rest are continuous. "Initial" refers to the first year. Payroll is in thousands of dollars. Panel 4 shows means of variables at the employee level. "Workers who join founding teams of startups ($\leq$3 yrs old)" refers to individuals who left the parent firm to join the startup's founding team (108,000 observations). "Workers who move to incumbents ($\geq$4 yrs old)" refers to individuals who left the parent firm to join already existing firms (2,900,000 observations). Wages are in thousands of real 2014 dollars.

Table 2: Effect of R&D on Non-entrepreneurial Employee Outcomes

Dependent variable:	Stayers$_{t+3}$		Depart employment data$_{t+3}$		Movers to incumbent firms$_{t+3}$	
	OLS	IV	OLS	IV	OLS	IV
	(1)	(2)	(3)	(4)	(5)	(6)
Firm log R&D$_{t-1}$	-1.13	-0.39	-0.004	0.46	0.49	-2.22
	(0.72)	(3.54)	(0.133)	(0.87)	(0.61)	(2.36)
Controls	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Firm FE	Yes	Yes	Yes	Yes	Yes	Yes
State FE	Yes	Yes	Yes	Yes	Yes	Yes
Industry (SIC3) FE	Yes	Yes	Yes	Yes	Yes	Yes
N	36,000	36,000	36,000	36,000	36,000	36,000
Adj. R^2	0.385	-	0.222	-	0.356	-

Dependent variable:	Movers to young firms (<10 yrs old)$_{t+3}$		Movers to high-tech young firms (<10 yrs old)$_{t+3}$	
	OLS	IV	OLS	IV
	(7)	(8)	(9)	(10)
Firm log R&D$_{t-1}$	0.35**	1.50*	0.23*	0.76
	(0.17)	(0.79)	(0.13)	(0.61)
Controls	Yes	Yes	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
Adj. R^2	0.294	-	0.253	-

Note: This table shows the effect of R&D on alternative employee outcomes. The sample is an establishment-year panel of public firms. In columns 1 and 2, the dependent variable is the percent of an establishment's workers in the 1st quarter of year zero who remain at the firm in the 1st quarter of year 3. In columns 3 and 4, the dependent variable is the percent of an establishment's workers in the 1st quarter of year zero who drop out of the employment sample by the 1st quarter of year three. Note, workers mainly depart the employment data coverage due to unemployment and departures out of the labor force (around 94 percent of workers in this group). Departures outside of the LEHD data coverage comprise around 6 percent of this group. In columns 5 and 6, the dependent variable is the percent of an establishment's workers in the 1st quarter of year zero who move to a firm that is more than 3 years old by the 1st quarter of year three. In columns 7 and 8, the dependent variable is the percent of an establishment's workers in the 1st quarter of year zero who move to a firm that is less than 10 years old by the 1st quarter of year three. In columns 9 and 10, the dependent variable is the percent of an establishment's workers in the 1st quarter of year zero who move to a firm that is 10 or more years old and also high-tech by the 1st quarter of year three. Controls are the same as in Table 3 Panel 1. Standard errors are clustered by firm. $*$, $**$, and $***$ denote significance at the 10%, 5%, and 1% levels.

Table 3: Effect of R&D on Employee Entrepreneurship

Panel 1

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

	(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)
Establishment log employment$_t$		-0.217***	-0.181***	-0.174***	-0.179***
		(0.018)	(0.019)	(0.018)	(0.019)
Establishment log payroll$_t$		-0.147***	-0.057	-0.082	-0.033
		(0.054)	(0.054)	(0.056)	(0.054)
Firm age$_{t-1}$		-0.036	-0.033	-0.021	-0.003
		(0.036)	(0.033)	(0.028)	(0.030)
Firm diversified$_{t-1}$		-0.123	-0.130	-0.135	-0.141
		(0.095)	(0.095)	(0.095)	(0.100)
Firm sales growth$_{t-1}$		0.126	0.130	0.124	0.129
		(0.089)	(0.090)	(0.091)	(0.099)
Firm ROA$_{t-1}$		0.131	0.127	0.155	-0.112
		(0.261)	(0.260)	(0.261)	(0.294)
Firm investment/total assets$_{t-1}$		0.888	0.811	0.731	0.508
		(0.543)	(0.543)	(0.553)	(0.617)
Firm log Tobin's Q$_{t-1}$		0.022	0.032	0.027	0.044
		(0.066)	(0.067)	(0.067)	(0.077)
Firm log total assets$_{t-1}$		-0.011	-0.033	-0.054	-0.001
		(0.070)	(0.069)	(0.070)	(0.066)
Firm PPE investment/total assets$_{t-1}$		-0.177	-0.058	-0.050	-0.063
		(0.382)	(0.385)	(0.393)	(0.424)
Firm cash/total assets$_{t-1}$		-0.526*	-0.502	-0.506	-0.521
		(0.308)	(0.307)	(0.315)	(0.320)
Firm leverage$_{t-1}$		-0.016	0.052	0.069	0.187
		(0.227)	(0.220)	(0.225)	(0.203)
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

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

	(1)	(2)	(3)
Firm log R&D$_{t-1}$	0.102**	0.104**	0.101**
	(0.052)	(0.051)	(0.051)
Establishment employee average age$_t$		-0.036***	
		(0.007)	
Establishment employee share female$_t$		-0.084	
		(0.165)	
Establishment employee share white$_t$		0.713***	
		(0.169)	
Establishment employee share foreign$_t$		0.508**	
		(0.251)	
Establishment employee average education$_t$		-0.055	
		(0.043)	
Establishment employee average tenure$_t$		-0.023*	
		(0.013)	
Establishment employee average experience$_t$		0.004	
		(0.017)	
Firm log patent classes$_{t-1}$			0.227*
			(0.120)
Firm log patents$_{t-1}$			-0.137
			(0.091)
Firm log forward citations$_{t-1}$			-0.006
			(0.022)
Firm log backward citations$_{t-1}$			-0.005
			(0.038)
Controls	Yes	Yes	Yes
Year FE	Yes	Yes	Yes
Firm FE	Yes	Yes	Yes
State FE		Yes	Yes
Industry (SIC3) FE		Yes	Yes
Industry (SIC4) FE	Yes		
N	36,000	36,000	36,000
Adj. R^2	0.181	0.179	0.176

Note: This table shows the effect of corporate R&D on employee entrepreneurship. The sample is an establishment-year panel of public firms. The dependent variable 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. In Panel 2, controls are the same as in Panel 1. Standard errors are clustered by firm. $*$, $**$, and $***$ denote significance at the 10%, 5%, and 1% levels.

Table 4: Effect of Alternative Measures of R&D on Employee Entrepreneurship

Dependent variable:	Employee entrepreneurship$_{t+3}$	# Employee-founded startups$_{t+3}$	Employee entrepreneurship$_{t+3}$		
	(1)	(2)	(3)	(4)	(5)
Firm has above median Δ R&D$_{t-1}$	0.089***	0.070***			
	(0.033)	(0.024)			
Firm has top 10 pct Δ R&D$_{t-1}$			0.132**		
			(0.067)		
Firm has bottom 10 pct Δ R&D$_{t-1}$				-0.105**	
				(0.053)	
Firm R&D/total assets$_{t-1}$					0.887*
					(0.529)
Controls	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes
Firm FE	Yes	Yes	Yes	Yes	Yes
Industry (SIC3) FE	Yes	Yes	Yes	Yes	Yes
State FE	Yes	Yes	Yes	Yes	Yes
N	36,000	36,000	36,000	36,000	36,000
Adj. R^2	0.176	0.154	0.176	0.176	0.180

Note: This table shows the effect of alternative measures of R&D on employee entrepreneurship. The sample is an establishment-year panel of public firms. Change (Δ) in R&D is defined as: $\frac{R\&D_{t-1}-R\&D_{t-2}}{.5\cdot(R\&D_{t-1}+R\&D_{t-2})}$. Top 10 pct Δ R&D$_{t-1}$ is 1 if the firm had a change in R&D that is in the top 10 percentiles in that year, and 0 if it is in the bottom 90 percentiles. Bottom 10 pct Δ R&D$_{t-1}$ is defined analogously. In columns 1, 3, 4 and 5, the dependent variable is the percent of an establishment's workers as of the first 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. In column 2, the the dependent variable is the number of unique startups associated with entrepreneurs in the column 1 definition normalized by the pre-period employment and expressed in percentage points. Controls are the same as in Table 3 Panel 1. Standard errors are clustered by firm. $*$, $**$, and $***$ denote significance at the 10%, 5%, and 1% levels.

Table 5: Instrumented Effect of R&D on Employee Entrepreneurship

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

	(1)	(2)	(3)	(4)	(5)	(6)
Instrumented firm log R&D$_{t-1}$	0.577***	0.719***	0.659**	0.648**	0.587*	0.598**
	(0.207)	(0.274)	(0.271)	(0.270)	(0.317)	(0.276)
Controls		Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes		Yes
Firm FE	Yes	Yes	Yes	Yes	Yes	Yes
State FE				Yes		Yes
Industry (SIC3) FE			Yes	Yes		
Industry (SIC3)-year FE					Yes	
State-year FE					Yes	
Industry (SIC4) FE						Yes
N	36,000	36,000	36,000	36,000	36,000	36,000

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 is shown in Table A.4. The dependent variable 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. We do not display controls because Census limits the number of coefficients we may disclose. Establishment controls are size and average wage. Firm controls are return on assets, sales growth, Tobin's Q, asset tangibility (PPE investment/total assets), size (log total assets), cash holdings, age, and an indicator for being diversified (having establishments in more than one SIC three-digit industry). Standard errors are clustered by firm. $*$, $**$, and $***$ denote significance at the 10%, 5%, and 1% levels.

Table 6: Effect of R&D on Alternative Measures of Employee Entrepreneurship

Panel 1

Dependent variable:	# Employee-founded startups$_{t+3}$		Employee entrepreneurship$_{t+3}$: Incorporated startups only		Employee entrepreneurship$_{t+3}$: Top 3 earners (rather than top 5)	
	OLS	IV	OLS	IV	OLS	IV
	(1)	(2)	(3)	(4)	(5)	(6)
Firm log R&D$_{t-1}$	0.067*	0.324	0.083**	0.464**	0.099**	0.507**
	(0.037)	(0.200)	(0.042)	(0.234)	(0.040)	(0.226)
Controls	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Firm FE	Yes	Yes	Yes	Yes	Yes	Yes
State FE	Yes	Yes	Yes	Yes	Yes	Yes
Industry (SIC3) FE	Yes	Yes	Yes	Yes	Yes	Yes
N	36,000	36,000	36,000	36,000	36,000	36,000
Adj. R^2	0.154	-	0.183	-	0.154	-

Panel 2

Dependent variable:	Employee entrepreneurship if present at startup founding$_{t+3}$		Employee entrepreneurship$_{t+2}$		Employee entrepreneurship$_{t+2}$: 1-yr old startups only	
	OLS	IV	OLS	IV	OLS	IV
	(1)	(2)	(3)	(4)	(5)	(6)
Firm log R&D$_{t-1}$	0.082*	0.595**	0.076*	0.472**	0.057*	0.175
	(0.048)	(0.254)	(0.042)	(0.240)	(0.033)	(0.166)
Controls	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Firm FE	Yes	Yes	Yes	Yes	Yes	Yes
State FE	Yes	Yes	Yes	Yes	Yes	Yes
Industry (SIC3) FE	Yes	Yes	Yes	Yes	Yes	Yes
N	36,000	36,000	36,000	36,000	36,000	36,000
Adj. R^2	0.171	-	0.131	-	0.097	-

Note: This table shows the effect of R&D on alternative measures of employee entrepreneurship. The sample is an establishment-year panel of public firms. For a detailed description of the dependent variables, see Section 5.4. In Panel 1 columns 4 and 5, we consider only the top three earners of a startup in identifying employee entrepreneurs rather than the top five. In Panel 2 columns 1 and 2, we require the employee entrepreneur to have been present at the startup founding, which is the first year the firm is observed in the LBD. Panel 2 columns 3-6 require the employee to have moved to a startup founded after R&D is measured within two years (rather than three), and columns 5 and 6 further restrict the startup sample to be those that are only one year old as of $t + 2$. Controls are the same as in Table 3 Panel 1. Standard errors are clustered by firm. $*$, $**$, and $***$ denote significance at the 10%, 5%, and 1% levels.

Table 7: Parent Variation in Effect of R&D on Employee Entrepreneurship

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

	(1)	(2)	(3)	(4)
Firm log R&D$_{t-1}$	0.048	0.099*	0.016	0.103**
	(0.057)	(0.052)	(0.062)	(0.052)
Firm log R&D$_{t-1}$·Establishment high tech industry	0.083***			
	(0.029)			
Firm log R&D$_{t-1}$·Firm high patent generality$_{t-1}$		0.027*		
		(0.016)		
Firm log R&D$_{t-1}$·Firm large$_{t-1}$			0.130**	
			(0.056)	
Firm log R&D$_{t-1}$·Firm IPO$_{t-3,t-1}$				0.072
				(0.057)
Controls	Yes	Yes	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
Adj. R^2	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. Large is 1 if the parent has above-median total assets (calculated at the firm-year level), and 0 if assets are below median. 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. IPO equals 1 if the firm went public in the past three years, and 0 otherwise. 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 the same as in Table 3 Panel 1. Standard errors are clustered by firm. $*$, $**$, and $***$ denote significance at the 10%, 5%, and 1% levels.

Table 8: Effect of R&D on Employee Entrepreneurship by Startup Characteristics

Panel 1: Predictors of VC-backed, employee-founded startups

Dependent variable: Employee-founded startup ever received VC

	(1)	...*Continued*	
Firm log R&D$_{t-1}$	0.007***		
	(0.001)		
Employee age$_t$	0.001**	Establishment log employment$_t$	0.001
	(0.000)		(0.001)
Employee age$^2{}_t$	-0.000**	Establishment average wage$_t$	0.012***
	(0.000)		(0.003)
Employee female	-0.013***	Firm age$_{t-1}$	-0.002***
	(0.002)		(0.001)
Employee white	0.003**	Firm diversified$_{t-1}$	-0.003
	(0.001)		(0.006)
Employee foreign	-0.002	Firm sales growth$_{t-1}$	0.004
	(0.004)		(0.006)
Employee born in state	-0.007***	Firm ROA$_{t-1}$	-0.008
	(0.001)		(0.016)
Employee education	0.001***	Firm investment/total assets$_{t-1}$	-0.013
	(0.000)		(0.041)
Employee experience$_t$	-0.000	Firm log Tobin's Q$_{t-1}$	0.002
	(0.001)		(0.004)
Employee tenure$_t$	-0.000	Firm log Total Assets$_{t-1}$	-0.006***
	(0.000)		(0.002)
Employee log wages$_t$	0.008***	Firm PPE investment/total assets$_{t-1}$	-0.004
	(0.002)		(0.012)
Employee-founded startup age$_{t+3}$	0.007***	Firm cash/total assets$_{t-1}$	0.076***
	(0.001)		(0.015)
Employee-founded startup initial emp.	0.008***	Firm leverage$_{t-1}$	0.009
	(0.002)		(0.007)
Continued...		Year-state FE	Yes
		Year-industry (SIC3) FE	Yes
		N	108,000
		Adj. R^2	0.079

Note: This table shows the effect of R&D on characteristics of employee entrepreneurship. The sample is at the employee-founded startup level. Based on the main variable used in Table 3, we identify whether the new firm associated with the departing employee has a given characteristic. The dependent variable in Panel 1 is 1 if the employee-founded startup ever received VC backing (either before or after the employee-founded startup is identified in year $t+3$), and 0 if not. The "Employee..." controls in Panel 1 column 1 refer to the employee who left the parent to found a new firm. Standard errors are clustered by parent firm. $*$, $**$, and $***$ denote significance at the 10%, 5%, and 1% levels.

Panel 2: Employee-founded startup characteristics

Dependent variable:	Employee-founded startup...					
	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)
Firm log $\text{R\&D}_{t-1}$	0.008***	0.009***	0.028***	0.007**	-0.007**	0.002
	(0.003)	(0.004)	(0.006)	(0.003)	(0.003)	(0.002)
Controls	Yes	Yes	Yes	Yes	Yes	Yes
Year-state FE	Yes	Yes	Yes	Yes	Yes	Yes
Year-industry FE	Yes	Yes	Yes	Yes	Yes	Yes
N	108,000	108,000	108,000	108,000	108,000	108,000
Adj. R^2	0.080	0.102	0.318	0.083	0.206	0.053

Note: This table shows the effect of R&D on characteristics of employee entrepreneurship. The sample is at the employee-founded startup level. Based on the main variable used in Table 3, we identify whether the new firm associated with the departing employee has a given characteristic. The dependent variable in Panel 2 column 1 (2) (4) (5) (6) is 1 if the employee-founded startup is an incorporated business (is in a high-tech industry) (the employee-founded startup exited, which is comprised primarily of firm failures but likely includes a small share of acquisitions, by year 5) (in the same two-digit SIC code as the parent establishment) (is in the same state as the parent establishment), and 0 if not. In column 3, the dependent variable is the departing employee entrepreneur's log wages at the new firm in the 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. Controls in Panel 2 are the same as in Panel 1, except that we include the indicator for being VC-backed as an additional control in Column 4. Standard errors are clustered by parent firm. $*$, $**$, and $***$ denote significance at the 10%, 5%, and 1% levels.

Panel 3: 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 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
Year-state FE	Yes	Yes	Yes	Yes
Year-industry (SIC3) FE	Yes	Yes	Yes	Yes
N	108,000	108,000	108,000	108,000
Adj. R^2	0.195	0.115	0.035	0.157

Note: This table shows the effect of R&D on employee entrepreneurship based on the supply chain relationship between the parent and the employee-founded startup. The sample is at the employee-founded startup level. 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). Controls are the same as in Table 8 Panel 1. Standard errors are clustered by parent firm. $*$, $**$, and $***$ denote significance at the 10%, 5%, and 1% levels.

Internet Appendix

A.1 Instrumental variables calculation and discussion

Following Bloom et al. (2013), we focus on the tax price component of the user cost of R&D capital. When this equals one, R&D is taxed like any other spending. When it is below one, there is a government-provided tax incentive to conduct R&D. There are two instruments: federal and state tax credit changes. These vary across firms for two reasons. First, the geographic distribution of a firm's R&D activity will determine its exposure to state R&D tax credits and corporation tax regimes. Like Bloom et al. (2013), we use the Wilson (2009) state R&D tax price histories together with inventor locations to approximate where the R&D occurs. These allow us to compute a firm's state R&D tax price. Second, also like Bloom et al. (2013), we follow Hall (1993) to calculate a firm-specific user cost based on the federal rules, which has a firm-specific component in part because the definition of expenditure that can be applied to the federal R&D tax credit depends on a firm-specific "base."

As with any instrument, causal interpretation rests on assumptions about an underlying economic model. One assumption is relevance: Firms must react to R&D tax credits by increasing their R&D investment. The underlying model is one where a lower cost of capital for R&D leads firms to spend more on R&D. We use changes in R&D to suggest that the overall level of corporate R&D may be an important source of high-growth startups. R&D tax credit changes are unlikely to yield a measurable effect on the state's aggregate number of startups if they affect startup creation only through marginal corporate R&D changes. It is important to emphasize that this paper is not about the macroeconomic effects of R&D tax credits. Instead, we are focused on whether corporate R&D yields labor reallocation, in general, and to startups, specifically.

A.1.1 Instrument motivation and assumptions

We use two instruments: federal and state tax credit changes, which affect the firm's cost of R&D capital. Changes in tax credits affect firm incentives to invest in R&D, because they change the firm-specific user cost of R&D capital. The tax credits are not deductions. Instead, they reduce the firm's corporate income tax liability by the value of the credit. As with any instrument, causal interpretation rests on assumptions about an underlying economic model (Kahn & Whited 2017). One assumption is relevance: Firms must react to R&D tax credits by increasing their R&D investment. The underlying model is one where a lower cost of capital for R&D leads firms to spend more on R&D. The literature has established that R&D tax credits have strong effects on corporate R&D in the short and long term. For the federal tax credit, the elasticity has been estimated to be well below -1, such that reducing the tax price of R&D capital by one dollar stimulates more than a dollar of additional R&D. This evidence includes Hall (1993), McCutchen (1993), Mamuneas & Nadiri (1996), Hall & Van Reenen (2000), Billings et al. (2001), Bloom et al. (2002), and Clausen (2009). The high sensitivity of expenditure to the R&D tax credit may reflect the fact that firms tend to finance R&D out of free cash flows (Brown & Petersen 2011).[40]

There is also evidence that state R&D tax credits increase R&D within the affected state, as shown by Paff (2005) and Wu (2008), among others. The most conservative finding is in Wilson (2009), where a one percentage point increase in the state tax credit rate increases R&D by 1.7 percent in the short term and 3 to 4 percent in the longer term. Wilson (2009) also finds that the tax credits lead firms to reallocate R&D geographically. Since large, multi-state firms are responsible for most R&D expenditures, and they may shift R&D across states in response to the tax credits while our independent variable is firm-wide R&D, we expect the state instrument to be weaker than the federal one. Appendix Table A.4 bears this out.

R&D stimulated by tax credits does not simply reflect relabeling of existing related expenditures. Dechezleprêtre et al. (2016) show that a UK R&D tax credit increases patenting and citations. Balsmeier, Kurakina & Fleming (2018) find that California's R&D tax credit increases patenting, and the additional patents are particularly valuable. Beyond patents, Czarnitzki et al. (2011) and Cappelen et al. (2012) find positive effects of tax credits on product and process innovation, respectively.

[40]There is similar evidence of large positive elasticities for foreign programs, including in Canada and the UK (Dechezleprêtre, Einiö, Martin, Nguyen & Van Reenen 2016, Agrawal et al. 2014, and Guceri & Liu 2017).

We discuss the exclusion restriction in the context of entrepreneurship, which is the only outcome for which we find an effect. The restriction is that tax credits cannot affect employee entrepreneurship except through the employer's R&D. Startups do not usually use R&D tax credits because they typically do not have taxable income (Bankman & Gilson 1999).[41] We demonstrate in Appendix A.1.4 that there is no relation between the state tax credits and startup creation. This is consistent with our main result. We use changes in R&D to suggest that the overall level of corporate R&D may be an important source of high-growth startups. R&D tax credit changes are unlikely to yield a measurable effect on the state's aggregate number of startups if they affect startup creation only through marginal corporate R&D changes. It is important to emphasize that this paper is not about the macroeconomic effects of R&D tax credits. Instead, we are focused on whether corporate R&D yields employee-founded startups.

A.1.2 The Federal R&D tax credit

The first instrument is the federal tax price of R&D, which we denote as ρ^F_{ft}. Implemented in 1981, the federal "Research and Experimentation" (R&E) tax credit permits firms to reduce their corporate income tax liability by the value of the credit. The credit was extremely complex to calculate (leading to a substantial simplification in 2009), and has changed over time. In the early 2000s, the total value of the federal credits was about $5 billion per year (Wilson et al. 2005). In this description, we focus on the calculation of the credit between 1990 and 2005, which is the period during which we seek to predict public firm R&D.[42]

The credit has annual changes for most firms and is firm-specific for a number of reasons, including the nonlinear relationship between current R&D and the "fixed base" that determines the amount of tax credit that can be claimed (Hall 1993). Our calculation of ${\rho_{ft}}^F$ closely follows Hall (1993) and Bloom et al. (2013), and we refer the reader to those articles for more detailed information. One key component is that the credit relies on the qualified research expenditure (QRE) in that year;

[41] The presence of carryforwards may make the credits somewhat useful to some startups, but our evidence in the Appendix suggests any effect is not large enough to be measurable.

[42] The calculation was quite different before 1989. In practice, we draw heavily from code originally written for Hall (1993).

in particular, on the difference between actual QRE and a firm-specific "base."[43] Firms defined in the law as "startups" have a different treatment. Their base is three percent, and the base then changes each year. Also, the credit lapsed in 1995-96, which generates additional within-firm variation. Note because of the nonlinearity, we can control directly for firm age and other variables in the IV. Appendix Table A.8 shows that firm R&D is robustly dependent on the tax price even after we include third polynomials of sales and firm age (columns 5-6). We find substantial variation in ${\rho_{ft}}^F$ within firm over time. Appendix Table A.8 also shows that the instrument remains strong after controlling for past R&D (columns 1-4). ${\rho_{ft}}^F$ can take values between zero and one. If ${\rho_{ft}}^F = 1$, then the firm should not treat R&D differently than other expenditures. If ${\rho_{ft}}^F < 1$, R&D is less expensive than other expenditures. A cheaper user cost of capital should incentivize the firm to invest more in R&D. As expected, we show in Appendix Table A.6 that when the tax price is below one, R&D increases as ${\rho_{ft}}^F$ decreases.

A.1.3 State R&D tax credits

Following Bloom et al. (2013), our state instrument uses inventor locations to approximate where the R&D occurs. The state instrument requires two objects: the state tax price component of the R&D user cost of capital ($\rho^S_{s,t}$), and a measure of the share of a firm's R&D that occurs in a given state ($\theta_{f,s,t}$). We use the state tax price of R&D from Wilson (2009), which incorporates state level corporate income taxes, depreciation allowances, and R&D tax credits. These credits vary across states and time. To build $\theta_{f,s,t}$, we follow Bloom et al. (2013) and use the share of the firm's patent inventors located in state s. The firm's state-level tax price is then $\rho^S_{f,t} = \sum_s \theta_{f,s,t}\rho^S_{s,t}$. The state tax credit must be aggregated to the firm-level because R&D is only observable at the firm level. That said, it is also intuitive because the outputs from R&D investment may depart the firm from a variety of establishment types, including headquarters, manufacturing plants, and laboratories. Where R&D is located for tax purposes need not be the place where R&D outputs depart the firm via employees.

The first state R&D tax credit was implemented in 1982 by Minnesota; by the end of our sample period, 40 states had some sort of R&D tax credit. The calculation of the base amount, and the

[43]For the legal definition of QRE, see here: https://www.law.cornell.edu/uscode/text/26/41.

definition of qualified R&D, can vary across states (Wilson et al. 2005). According to Miller & Richard (2010), manufacturing-intensive states, and those with one-party political control, are more likely to pass R&D tax credits. They argue that the tax credits primarily support incumbent R&D-conducting firms. To the best of our knowledge, the state credits are not refundable during the sample period.

The state instrument requires two objects: the state tax price component of the R&D user cost of capital, and a measure of the share of a firm's R&D that occurs in a given state. For both, we follow Bloom et al. (2013). First, we use the state tax price of R&D in Wilson (2009). He incorporated state level corporate income taxes, depreciation allowances, and R&D tax credits into this tax price component, which we call ρ^S_{st}.[44] These credits vary across states and time. They allow a firm to offset its state-level corporate tax liabilities, and they are calculated by weighting total firm profits according to the location of the firm's sales, employment, and property. Thus firms with R&D activities in the state will likely both have tax liability and R&D tax credit eligibility there.

The second object, θ_{fst}, is a proxy for a firm's R&D share in a given state-year. It is the 10-year moving average of the share of the firm's patent inventors located in state s.[45] The firm's state-level tax price is then $\rho^S_{ft} = \sum_s \theta_{fst}\rho^S_{st}$.

A.1.4 Concerns

There are five potential concerns. Most importantly, the exclusion restriction is that tax credits cannot affect employee entrepreneurship. In a rigorous border county differences-in-differences model, Curtis & Decker (2018) show that R&D tax credits have no effect on startup formation. Wilson (2009) and Lucking (2018) also do not find employment effects among new firms, only existing firms. We also show empirically that there is no relation between the state tax credits and state-level startup creation, or the federal tax credit and national startup creation. We do this using two data sources, each of which have limitations. The first is the Business Dynamics Statistics (BDS), which contains firm entry by state for our entire sample period, but does not have state-industry data.[46] The second is the Quarterly

[44] Specifically, it is roughly: $\frac{1-(tax\ credits+depr.\ allowances)}{1-tax\ rate}$.
[45] The data is from NBER patent data, available here.
[46] This public version of the LBD is available here.

Workforce Indicators (QWI), a publicly available dataset derived from the LEHD. While the QWI has state-industry level data, its coverage is poor in the early years of our data, with states being added over time.[47]

At the state level we regress either the log number of new firms or the change in firm entry rates year to year on the tax price of R&D, as well as state and year fixed effects. The results are in Appendix Table A.9 Panels 1-2. We consider only R&D-intensive industries and cluster errors by state.[48] We find that the tax credits have no correlation with startup entry (BDS data) or new firm employment growth (QWI data). This is true regardless of whether we use year and/or state fixed effects, and regardless of the standard error assumptions. At the federal level, we regress either the log number of new firms or the change in firm entry rates on the statutory federal R&D tax credit. This is, of course, very different from the firm-specific tax price of R&D that is calculated per the description in Section 7. This reflects baseline changes in the rate, which is then applied to a firm's specific situation. There are very few observations, and we do not use robust standard errors. The results, in Appendix Table A.9 Panel 3, again show no correlation.

More generally, the legal literature has argued that R&D tax credits are not useful to startups, because these firms have no or little taxable income against which to offset losses from failed R&D efforts. For example Bankman & Gilson (1999) note that "the U.S. tax code subsidizes R&D by existing successful companies by allowing losses from failed attempts at innovation to offset otherwise taxable income from other activities. Since startups have no other income against which their losses from a particular project may be set off, the government in effect gives established companies with a stable source of income an R&D tax subsidy that is not available to a startup entity." Perhaps in response to this, a few states have recently made their R&D tax credits transferable, so that firms without revenue can potentially derive value from them. However, these policies occurred after the end of our sample period.

The second concern is that changes in state-level R&D tax credits may lead firms to geographically reallocate R&D (or misreport it such that it appears reallocated). For studies

[47] We used a transformed version of the QWI data courtesy of Song Ma.
[48] NAICS codes 31-33, 51, and 54.

evaluating how a state-level R&D tax credit affects national R&D, this is a central concern. In our case, however, such reallocation will simply reduce the power of the instrument. As long as the combined instruments have adequate power, some degree of reallocation should not bias our findings. It does lead us to expect that the federal instrument will have more power than the state instrument, which is indeed what we find. This is because the federal instrument should have a larger effect on firms that only operate in the affected state, but most firms with positive R&D operate in multiple states.

The third concern is that the tax credits may not be large enough to affect R&D. The above sections pointed to substantial literature finding R&D responses to R&D tax credits that are large in economic magnitude and quite robust, especially for the federal instrument. The literature examining the state instrument finds large within-state elasticities, but also finds evidence of reallocation across states.

The fourth concern is that changes to the R&D tax credits may be anticipated by firms, which may then behave strategically to maximize their value. The federal tax credit formula is exceedingly complicated, as explained above, and it seems implausible that firms will optimize on all of the variables (especially firm age) in order to maximize the tax credit value. Strategic behavior around state tax credit changes would require firms in one state to respond by moving states. The tax credits are not large enough to merit such a response from many firms. For firms in multiple states, reallocation across states should attenuate the effect of the instrument. Beyond these points, note that the goal is to predict changes in R&D. Suppose that in order to maximize the tax credit benefit, firms choose to conduct less R&D in the years immediately preceding the tax credit change and more in the years after. This does not obviously bias our main result, which is that changes to R&D affect employee entrepreneurship.

Finally, the fifth concern is that state decisions to adopt R&D tax credits could be endogenous, reflecting recent declines in R&D. Bloom et al. (2013) consider this possibility at length, and show that the results are robust to lagging the tax credit instruments for one and two periods. They also point out that cross-sectional variation in the state R&D tax credit rates is very large relative to the average rate within states, and also large relative to the secular increase in the tax credit generosity that has occurred over time. Finally, Chirinko & Wilson (2008), Chirinko & Wilson (2011), and Bloom et al. (2013) show that the level and timing of R&D tax credit adoption is uncorrelated with local economic observables like state R&D expenditure or per capita GDP, once year and state fixed effects are included.

A.1.5 First Stage Estimation

In sum, we believe that R&D tax credits offer the best available quasi-exogenous source of variation driving corporate R&D that is plausibly unrelated to technological opportunities that could jointly give rise to parent R&D and employee entrepreneurship. Most importantly, they are plausibly unrelated to technological or demand shocks that could jointly give rise to parent R&D and employee reallocation. Having constructed the firm-level federal and state tax prices of R&D ($\rho^F_{f,t}$ and $\rho^S_{f,t}$ respectively), we estimate the following first stage regression:

$$\begin{aligned} \ln(R\&D_{f,t}) = \beta_1 \ln\left(\rho^S_{f,t}\right) + \beta_2 \ln\left(\rho^F_{f,t}\right) + \text{Firm FE}_f + \text{Industry-year FE}_{e,t} \\ + \text{State-year FE}_{e,t} + \xi \text{Controls}_{f,t} + \zeta \text{Controls}_{e,t} + \varepsilon_{e,f,t}. \end{aligned} \tag{3}$$

R&D tax credits could have other effects in the state. 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.4. Despite the different time period, they are roughly similar in magnitude to the first stage estimates in Bloom et al. (2013). The instruments are strong, yielding F-statistics of about 25, well above the rule-of-thumb cutoff of 10. The partial R^2 of the two instruments ranges from 2.2 to 3.2 percent, which captures a reasonable amount of variation in R&D (Jiang 2015). Bloom et al. (2013) use only firm and year fixed effects. This is equivalent to column 1. Our preferred specification, with SIC three-digit industry-year and state-year fixed effects, along with firm time-varying controls and firm fixed effects, is in column 5.[49]

[49]We find that R&D tax credits do not predict total investment, only R&D investment. We are grateful to Shai Bernstein for suggesting this placebo test.

A.2 Out-of-sample test for benefit internalization

We directly assess the possibility that parents internalize employee-founded startups' benefits using an out-of-sample test based on the underlying data in Gompers et al. (2005). They connected all venture capital-backed startup executives in the VentureOne database to their prior employers.[50] This sample should give an upper bound on possible internalized employee-founded startups because as these startups by definition received external investment, they are more likely than the average employee-founded startup to have received investment from their former employer. We begin with 13,612 entrepreneur-parent pairs. The entrepreneurs are founders of 6,499 unique startups. There are 9,152 unique parents, which we linked to VentureXpert acquisition and investment data.[51] Seventy-four percent of the unique startups matched to at least one investor or acquirer, yielding 20,478 unique startup-investor pairs.[52]

Finally, among the unique investors and acquirers in these pairs, only 208 match to parents. This is just 2.3 percent of the 9,152 unique parents in the original Gompers et al. (2005) data, providing evidence that parents do not usually internalize employee-founded startups by investing in or acquiring them. There are 266 unique startups where the parent matches an investor or acquirer, 5.6 percent of the startups matched to VentureXpert.[53] Of these, 192 are investment deals, and 74 are acquisitions. Some parents have multiple startups, such as IBM and Highland Capital Partners, so the parent and startup numbers do not match. Some parents that invested in or acquired their employee-founded startups are corporates, including Seagate, Xerox, Monsanto, Johnson & Johnson, and Microsoft. Others are asset

[50] Their time period, 1986 to 1999, overlaps with our primary Census data (1990 to 2005).

[51] In many cases, employee-founded startups have multiple parents (that is, there are multiple executives with prior jobs).

[52] Note that the underlying dataset, from Dow Jones Venture Source, is of venture capital-backed startups. In theory, if we used VentureSource, we should match 100 percent to initial investors. However, as Kaplan & Lerner (2016) and Maats et al. (2011) explain, VentureXpert's coverage is much better than Venture Source (it contains more than 40 percent more investments). VentureXpert also has superior acquisition data, and Venture Source's data quality has declined over time. We are most interested in whether parents ultimately invested in (and especially acquired) employee-founded startups, so VentureXpert seems like the optimal data set to use. If there is any bias, it should be the case that the employee-founded startups that do not match have lower rates of subsequent investment and acquisition, since the commercial databases often backfill based on exit events.

[53] We matched on the company's first word, which yielded 275 matches. This enables successful matches such as "Xerox Venture Capital" to "Xerox." We then manually removed obviously wrong matches, erring on the side of leaving the match to be conservative in ambiguous cases.

managers, including Accel Partners, SoftBank Vision Group, and Equus Capital. Still others are non-corporates, including Boston University. We identified 41 parent firms that are clearly venture funds or other asset managers. This leaves 167 parents that are potentially R&D-investing companies.

One concern may be that many corporate parents may not be covered as investors or acquirers in VentureXpert. We match 2,617 of the parents to investors or acquirers in VentureXpert. The most conservative framing of our results, then, restricts the parent population to firms that ever invested in or acquired a startup in VentureXpert. In this case, 7.9 percent of parents (208 out of 2,617) invest in or acquire their employee-founded startups. This extreme upper bound is still small and confirms that it is unlikely that parents generally internalize the benefits of their employee-founded startups.

The parent could also appropriate the employee-founded startup's benefits through technology licensing deals. We cannot assess this possibility with our data, but we think it unlikely that the parent can fully internalize the employee-founded startup's social benefits through such arms-length contracts.

Consistent with the out-of-sample test, in our data we find that the interaction between R&D and the parent having a corporate venture capital program has no effect on employee entrepreneurship. These results are consistent with Ma (2016), who finds that public firms launch corporate venture capital programs when internal innovation is poor, invest in startups in their own industries, and invest in geographically distant startups. That is, corporate venture capital is a way to outsource innovation. This is the opposite of the corporate environment that yields R&D-induced employee entrepreneurship. Instead, when corporate R&D increases at innovative firms, it seems to serendipitously produce "extra" growth options, and employee entrepreneurship is an unintended consequence.

A.3 Benchmarking attrition with the CPS

The goal of this appendix is to better understand the expected degree of worker attrition from private employment over time. This information helps us gauge whether the attrition rates that we observe in the LEHD data should be considered as abnormally high, low, or aligned with expectations relative to the overall employment population.

To set the context, in our LEHD data, on average 12% of the workers observed in the LEHD data each year are not again observed in the LEHD data three years later. The LEHD data primarily samples workers in private companies and, as such, the likely main drivers of attrition are when workers: 1) move out of the labor force (e.g., retirement, maternity leave), 2) become unemployed, and/or 3) move to public sector employment not covered by the LEHD.

To validate the use of the LEHD in a longitudinal manner and address any concerns regarding this rate of attrition, we compare the rate in our sample to a benchmark rate of attrition calculated using out-of-sample data. The Current Population Survey (CPS) is the most appropriate U.S. source for developing an out of sample benchmark. First, the CPS is the current standard created by the U.S. Census Bureau and BLS for nationally representative employment statistics.[54] Second, the CPS is longitudinal in nature, allowing us to measure employee attrition over time. The CPS surveys a household for a maximum of a 16-month time interval, allowing us to measure the attrition for at most over 15 months.

We find that within the CPS data, after 12 (15) months 9.4% (9.9%) of the original private sector workers are no longer employed (become unemployed or move out of labor force). Comparing this CPS's *1-year* attrition rate of 9.4% to the *3-year* attrition rate of 12% of the workers in our sample indicates that our attrition rates are not abnormally high relative to a representative sample of U.S. workers. In addition, our coverage of the LEHD is limited to 31 states, thus moving into states not included in our sample is another driver of this attrition. However, we do not think this is a big reason for attrition in our data given that our data cover most of U.S. employment (about 60%) and out-of-state mobility over our time period is relatively low (1.5-2% of worker move across state lines per year).

[54]For example, CPS statistics underlie the monthly "Employment Situation" report that is the U.S. government's primary estimate of employment (https://stats.bls.gov/news.release/empsit.toc.htm).

Table A.1: Variable Definitions and Sources

Variable name	Definition	Level of observation	Source
R&D$_{t-1}$	Parent firm R&D expenditure as reported in 10-K filings	Firm-year	Compustat
R&D/Total Assets$_{t-1}$	Parent firm R&D divided by total assets	Firm-year	Compustat
Above median Δ R&D$_{t-1}$	Equals 1 if the parent firm has above-median total R&D (calculated at the firm-year level), and 0 if R&D is below median	Firm-year	Compustat
Top 10 pct Δ R&D$_{t-1}$	Equals 1 if the parent firm has R&D in the top 10 percent of the distribution (calculated at the firm-year level), and 0 if R&D is below the top 10 percent	Firm-year	Compustat
Bottom 10 pct Δ R&D$_{t-1}$	Equals 1 if the parent firm has R&D in the bottom 10 percent of the distribution (calculated at the firm-year level), and 0 if R&D is above the bottom 10 percent	Firm-year	Compustat
Tobin's Q$_{t-1}$	Equals to the firm market value of assets (market value of common equity plus total assets minus book value of common equity) divided by book value of total firm assets. The market value of common equity is the number of common shares times stock price at the end of the fiscal year. The book value of common equity is common equity plus deferred taxes and investment tax credit	Firm-year	Compustat
Diversified$_{t-1}$	Equals 1 if the parent firm has establishments in more than one three-digit SIC industry, and 0 otherwise	Firm-year	LBD
Firm Age$_{t-1}$	Years since the oldest establishment that the firm owns is first observed	Firm-year	LBD
Sales growth$_{t-1}$	Is the parent firm annual sales growth	Firm-year	Compustat
Total Assets$_{t-1}$ ('000s)	Is the parent firm total assets as reported in 10-K filings expressed in thousand of dollars	Firm-year	Compustat
Large$_{t-1}$	Equals 1 if the parent has above-median total assets (calculated at the firm-year level), and 0 if assets are below median, and 0 otherwise	Firm-year	Compustat
IPO$_{t-3,t-1}$	Equals 1 if the parent firm went public in the past three years, and 0 otherwise	Firm-year	Compustat
ROA$_{t-1}$	Is the parent firm operating income (EDITDA) normalized by total assets	Firm-year	Compustat
Investment/Total assets$_{t-1}$	Is the parent firm CAPEX minus sales of property, plant, and equipment divided by total assets	Firm-year	Compustat

Table A.1 Continued

Variable name	Definition	Level of observation	Source
PPE investment/Total assets$_{t-1}$	Is the parent firm property, plant, and equipment (PP&E) normalized by total assets	Firm-year	Compustat
Cash/Total Assets$_{t-1}$	Is the parent firm cash and short-term investments divided by total assets	Firm-year	Compustat
Payroll ('000s)	Is the parent firm total payroll in thousands of dollars calculated as the sum of payroll across all its establishments	Firm-year	LBD
Establishment Payroll ('000s)	Is the parent firm payroll in thousands of dollars at the establishment where the employee works	Establishment-year	LBD
Cash/Total Assets$_{t-1}$	Is the parent firm cash and short-term investments divided by total assets	Firm-year	Compustat
Leverage$_{t-1}$	Is the parent firm book financial leverage ratio (long-term debt plus debt in current liabilities, normalized by total firm assets). It is set to missing if negative and set to 1 if larger than 1	Firm-year	Compustat
Employment$_{t-1}$	Is the parent firm total employment calculated as the sum of employment across all its establishments	Firm-year	LBD
Patent classes$_{t-1}$	Is the number of patent classes a firm or industry patents in	Firm-year	NBER Patent Data
Patents$_{t-1}$	Is the number of granted patents applied for in a given year	Firm-year	NBER Patent Data
Had ≥ 1 patent$_{t-10,t}$	Equals 1 if the parent firm had at least one patent over the past 10 years, and 0 otherwise	Firm-year	NBER Patent Data
Forward citations$_{t-1}$	Is the number of citations made to this patent by future patents	Firm-year	NBER Patent Data
Backward citations$_{t-1}$	Is the number of citations in this patent to preexisting patents	Firm-year	NBER Patent Data
High patent generality$_{t-1}$	Equals 1 if the parent has above-median patent generality (calculated at the industry-year level), and 0 if it is below median. Generality is higher when forward citations are in many classes	Firm-year	NBER Patent Data
Made corporate VC investments$_t$	Equals 1 if the parent made corporate VC investments, based on matching corporate venture investors from CB Insights to Compustat	Firm-year	CB Insights
Federal R&D tax price$_{t-1}$	Is the log firm-level tax price of R&D, based on the federal tax credit	Firm-year	Hall (1993) and Bloom et al. (2013)

Table A.1 Continued

Variable name	Definition	Level of observation	Source
State R&D tax price$_{t-1}$	Is the log state-level tax price of R&D, based on the state tax credit	Firm-year	Wilson (2009)and Bloom et al. (2013)
Weak non-compete enforcement (state)	Equals 1 for establishments in states with the Noncompetition Enforceability Index value less than (greater or equal to) 5, the median value of the index, and 0 otherwise. The index measures how strictly states enforce non-compete agreements. The index is available for states from 1992 through 2004. For 1990 and 1991, we back-fill the values of the index from 1992. For 2015, we set the value equal to that in 2014	Establishment-year	Garmaise (2011)
Employee entrepreneurship$_{t+3}$	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. This is the main definition of entrepreneurship used in the paper	Establishment-year	LBD/LEHD
# employee-founded startups$_{t+3}$	Is the number of unique new firms associated with the founding team members defined according to the main definition. The number of firms is normalized by an establishment's employment as of 1st quarter of year zero	Establishment-year	LBD/LEHD
Stayers$_{t+3}$ (%)	Is the percent of an establishment's workers as of 1st quarter of year zero who are still observed at the R&D-investing parent firm as of the first quarter of year three	Establishment-year	LBD/LEHD
Movers to incumbent firms ($\geq$4 yrs old)$_{t+3}$	Is the percent of an establishment's workers as of 1st quarter of year zero who are observed at a different firm aged four years or more as of the first quarter of year three	Establishment-year	LBD/LEHD
Movers to new firms ($\leq$3 yrs old)$_{t+3}$	Is the percent of an establishment's workers as of 1st quarter of year zero who are observed at a different firm aged three years or fewer as of the first quarter of year three	Establishment-year	LBD/LEHD
Movers to young firms (<10 yrs old)$_{t+3}$	Is the percent of an establishment's workers as of 1st quarter of year zero who are observed at a different firm aged fewer than ten years as of the first quarter of year three	Establishment-year	LBD/LEHD
Movers to high-tech young firms (<10 yrs old)$_{t+3}$	Is the percent of an establishment's workers as of 1st quarter of year zero who, as of the first quarter of year three, are observed at a different firm which is aged fewer than ten years old and is also in high-tech industry	Establishment-year	LBD/LEHD

Table A.1 Continued

Variable name	Definition	Level of observation	Source
Depart employment data$_{t+3}$	Is the percent of an establishment's workers as of 1st quarter of year zero who are not observed in the LEHD data as of the first quarter of year three. According to the BLS data, workers depart employment mainly because they exit labor force. Other major departure reasons in our data are unemployment and geographic reallocation outside of the states covered by the LEHD. Using the BLS estimates on cross-state mobility and the percentage of labor force covered by our LEHD states, we estimate that only around six percent of those who exit coverage depart to locations outside our LEHD data coverage	Establishment-year	LBD/LEHD
Number of employees$_t$	Is the number of an establishment's (i.e., SEIN) workers plus one, measured as of the first quarter of the base year to reflect characteristics of the workers in the establishment-worker panel	Establishment-year	LBD/LEHD
Incorporated	Equals 1 if the new firm is an incorporated business, and 0 otherwise	Employee-founded startup	LBD
High-tech industry	For a new firm, it equals 1 if the firm is in the high-tech industry, and 0 otherwise. For an R&D-investing firm, it equals 1 if its establishment is in the high-tech industry, and 0 otherwise. High-tech industries include biotech, chemicals, software, business services, and high-tech manufacturing & R&D	Employee-founded startup; establishment-year	LBD
Startup exit in 5 years	Equals 1 if a start-up exits by the fifth year since it became identified with the founding team member, and 0 otherwise	Employee-founded startup	LBD
Ever received VC	Equals 1 for a firm having ever received venture capital, and 0 otherwise	Employee-founded startup	LBD,Puri & Zarutskie (2012)
Startup initial employment	Is the number of employees measured during the first year the new firm appears in the LBD with positive employment	Employee-founded startup	LBD
Startup initial payroll ('000s)	Is total payroll measured during the first year the new firm appears in the LBD with positive employment	Employee-founded startup	LBD
Employee female$_t$	Equals 1 if a worker is female, and 0 otherwise	Employee-year	LEHD

Table A.1 Continued

Variable name	Definition	Level of observation	Source
Employee white$_t$	Equals 1 if a worker is white, and 0 otherwise	Employee-year	LEHD
Employee foreign$_t$	Equals to 1 if a worker was born outside of the US, and 0 otherwise	Employee-year	LEHD
Employee education$_t$	Is the number of years of education. Note: education is imputed for some workers in the LEHD database	Employee-year	LEHD
Employee Age$_t$	Is the worker age in years.	Employee-year	LEHD
Employee tenure$_t$	Is the number of years a worker is at the establishment of a public firm. Note: since the LEHD coverage starts in 1990, the variable is left censored	Employee-year	LEHD
Wages ('000s)$_t$	Is the real wages earned at a public firm during the quarter the worker is identified with that firm. Real earnings are in constant 2014 dollars	Employee-year	LEHD
Employee experience$_t$	Is the number of years a worker is in the LEHD. Note: since LEHD coverage starts in 1990, the variable is left censored	Employee-year	LEHD
Employee born in state	Equals 1 if a worker was born outside the state of location of the establishment of a public firm, and 0 otherwise	Employee	LEHD
Employee-founded startup age$_{t+3}$	Is the new firm age measured as of the first quarter of year 3 when founding team members are identified with new firms	Employee-founded startup	LBD
Log wages$_{t+3}$	Is the log wages at the new firm of a departing employee who are identified as a founding team member at the startup in the 1st quarter of year three. A founding team member is a person at a firm no more than 3 years old who is among the top 5 earners at that new firm	Employee	LEHD
Same industry (SIC2) as parent	Equals 1 if the employee-founded startup is in the same two-digit SIC code as the parent establishment, and 0 otherwise	Startup-parent establishment pair	LBD
Same state as parent	Equals 1 if the employee-founded startup is in the same state as the parent establishment, and 0 otherwise	Startup-parent establishment pair	LBD

Table A.2: Sample Composition by Industry

Panel 1: 1990 -2001

	In Sample	Out of Sample
Construction	4.8%	4.1%
Finance, Insurance, and Real Estate	5.6%	6.3%
Manufacturing	15.4%	15.8%
Mining	0.6%	0.4%
Services	27.9%	28.8%
Total Government	16.4%	17.2%
Trade	23.7%	22.6%
Transportation and Public Utilities	5.5%	4.8%

Panel 2: 2002-2008

	In Sample	Out of Sample
Construction	5.6%	4.8%
Educational Services	1.9%	2.4%
Financial Activities	5.9%	6.3%
Government	16.3%	17.0%
Health Care and Social Assistance	10.9%	11.6%
Information	2.2%	2.5%
Leisure and Hospitality	9.9%	9.1%
Manufacturing	10.6%	10.7%
Mining and Logging	0.6%	0.3%
Other Services	4.0%	3.9%
Professional and Business Services	12.3%	12.8%
Retail Trade	11.6%	11.1%
Transportation and Warehousing	3.4%	2.8%
Utilities	0.4%	0.4%
Wholesale Trade	4.4%	4.2%

Note: This table compares the data in our sample (from 31 states) to national data from the Bureau of Labor Statistics (BLS) Current Employment Statistics (CES) Survey from 1990-2008. This is done separately for the pre-2002 and post-2002 periods because before 2002, the BLS used SIC codes, while after 2002, it used NAICS codes. Panel 1 shows the pre-2002 industries, and Panel 2 the post-2002 industries. We divide state-industry level employment by total state employment across all states in our sample. We do this for each year, and then average across years. We compare this to the analogous figure for states that are not in our sample (right column).

Table A.3: Industry Composition by Sample

Panel 1: 1990 - 2001

	In Sample	Out of Sample
Construction	63.7%	36.3%
Finance, Insurance, and Real Estate	57.4%	42.6%
Manufacturing	59.4%	40.6%
Mining	69.4%	30.6%
Services	59.2%	40.8%
Total Government	58.8%	41.2%
Trade	61.1%	38.9%
Transportation and Public Utilities	63.2%	36.8%
Total Observations	60.0%	40.0%

Panel 2: 2002-2008

	In Sample	Out of Sample
Construction	64.5%	35.5%
Educational Services	55.0%	45.0%
Financial Activities	59.6%	40.4%
Government	59.9%	40.1%
Health Care and Social Assistance	59.5%	40.5%
Information	57.5%	42.5%
Leisure and Hospitality	62.7%	37.3%
Manufacturing	60.6%	39.4%
Mining and Logging	71.9%	28.1%
Other Services	61.6%	38.4%
Professional and Business Services	59.9%	40.1%
Retail Trade	61.8%	38.2%
Transportation and Warehousing	65.3%	34.7%
Utilities	59.6%	40.4%
Wholesale Trade	62.3%	37.7%
Total Observations	62.3%	37.7%

Note: This table compares the data in our sample (from 31 states) to national data from the BLS Current Employment Statistics (CES) Survey from 1990-2008. This is done separately for the pre-2002 and post-2002 periods because before 2002, the BLS used SIC codes, while after 2002, it used NAICS codes. Panel 1 shows the pre-2002 industries, and Panel 2 the post-2002 industries. Each percent is the share of people employed in an industry in our sample states (left column) versus the other states (right column).

Table A.4: First Stage IV Results

Dependent variable: Firm log R&D$_{t-1}$

	(1)	(2)	(3)	(4)	(5)	(6)
Firm log federal R&D tax price$_{t-1}$	-2.020***	-1.504***	-1.504***	-1.470***	-1.363***	-1.424***
	(0.295)	(0.231)	(0.231)	(0.225)	(0.168)	(0.199)
Firm log state R&D tax price$_{t-1}$	-1.158*	-0.950**	-0.956**	-0.978**	-0.303	-0.947**
	(0.691)	(0.476)	(0.476)	(0.471)	(0.375)	(0.420)
Controls		Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes		Yes
Firm FE	Yes	Yes	Yes	Yes	Yes	Yes
State FE			Yes	Yes		Yes
Industry (SIC3) FE				Yes		
Industry (SIC3)-year FE					Yes	
State-year FE					Yes	
Industry (SIC4) FE						Yes
N	36,000	36,000	36,000	36,000	36,000	36,000
R^2 (partial for the IV instruments)	0.032	0.027	0.026	0.026	0.022	0.025
F-test (instruments)	24.70	22.23	22.25	22.37	34.11	27.64

Note: This table shows the first stage of the instrumental variables analysis. The sample is an establishment-year panel of public firms. We predict parent firm R&D using firm-level federal and state tax prices of R&D, which are partially determined by tax credits that change across time, states, and depending on firm age. The federal R&D tax price is the log firm-level tax price of R&D, based on the federal tax credit, and following Hall (1993) and Bloom et al. (2013). The state R&D tax price is the log state-level tax price of R&D, following Bloom et al. (2013). See Section 4.2 and Appendix Section A.1 for details. Establishment controls are size and average wage. Firm controls are return on assets, sales growth, Tobin's Q, asset tangibility (PPE investment/total assets), size (log total assets), cash holdings, age, and an indicator for being diversified (having establishments in more than one SIC three-digit industry). Standard errors are clustered by firm. $*$, $**$, and $***$ denote significance at the 10%, 5%, and 1% levels.

Table A.5: Effect of R&D on Employee Entrepreneurship at the Employee Level

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

	OLS	IV		OLS	IV	
		First stage	Second stage		First stage	Second stage
	(1)	(2)	(3)	(4)	(5)	(6)
Firm log R&D$_{t-1}$	0.00048**		0.003**	0.00039**		0.003**
	(0.00)		(0.00)	(0.00)		(0.00)
Firm log federal R&D tax price$_{t-1}$		-1.429***			-1.488***	
		(0.21)			(0.20)	
Firm log state R&D tax price$_{t-1}$		-0.152			-0.233	
		(0.56)			(0.53)	
Employee controls	Yes	Yes	Yes	Yes	Yes	Yes
Firm FE	Yes	Yes	Yes	Yes	Yes	Yes
State FE	Yes	Yes	Yes			
State-year FE				Yes	Yes	Yes
Industry (SIC3)-year FE	Yes	Yes	Yes	Yes	Yes	Yes
N (millions)	11.9	11.9	11.9	11.9	11.9	11.9
Adj. R^2	0.00			0.01		
R^2 (partial for the IV instruments)		0.02			0.02	
F-test (instruments)		23.10			26.80	

Note: This table shows the effect of R&D on employee entrepreneurship. The sample is an establishment-year panel of public firms. The dependent variable equals 1 if an employee as of first quarter of year zero becomes an entrepreneur as of 1st quarter of year three, and 0 otherwise. 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. The mean of the dependent variable is 0.0091. We do not display controls because we are limited by Census in the number of coefficients we may disclose. Employee controls are age, age squared, education, total experience in years, tenure at the firm, log earnings in year 0, and indicators for being female, white, foreign, and born in state. Standard errors are clustered by firm. $*$, $**$, and $***$ denote significance at the 10%, 5%, and 1% levels.

Table A.6: Reverse Causality Test (Effect of Employee Entrepreneurship on R&D)

Dependent variable: Firm log R&D_t

	(1)	(2)	(3)
One-year employee entrepreneurship$_{t-2,\,t-1}$	0.008		
	(0.005)		
Two-year employee entrepreneurship$_{t-3,\,t-1}$		0.001	
		(0.006)	
Three-year employee entrepreneurship$_{t-4,\,t-1}$			-0.005
			(0.003)
Controls	Yes	Yes	Yes
Year FE	Yes	Yes	Yes
Firm FE	Yes	Yes	Yes
State FE	Yes	Yes	Yes
Industry (SIC3) FE	Yes	Yes	Yes
N	36,000	36,000	36,000
Adj. R^2	0.879	0.879	0.879

Note: This table shows that past employee entrepreneurship does not predict current corporate R&D. The sample is an establishment-year panel of public firms. The independent variables are lagged variations on our main employee entrepreneurship rate measures used as the dependent variable in Tables 3 and 5. The one-year employee entrepreneurship$_{t-1}$ rate is the percent of an establishment's workers as of the first quarter of year $t-1$ who are entrepreneurs as of the first quarter of year t, which is the year that R&D is measured (the dependent variable). The two-year employee entrepreneurship$_{t-2}$ rate is the percent of an establishment's workers as of the first quarter of year $t-2$ who are entrepreneurs as of the first quarter of year t. The three-year employee entrepreneurship$_{t-3}$ rate is the percent of an establishment's workers as of the first quarter of year $t-3$ who are entrepreneurs as of the first quarter of year t. 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. Controls are the same as in Table 3 Panel 1. Standard errors are clustered by firm. $*$, $**$, and $***$ denote significance at the 10%, 5%, and 1% levels.

Table A.7: Relationship Between Federal Tax Price Regions and R&D

Dependent variable: Firm log R&D$_t$			
	(1)	(2)	(3)
$1 \mid$Firm federal R&D tax price$_t < 0.99$	0.18***		
	(0.016)		
$1 \mid 0.95 \leq$Firm federal R&D tax price$_t < 0.99$		0.15***	0.15***
		(0.016)	(0.017)
$1 \mid$Firm federal R&D tax price$_t < 0.95$		0.22***	
		(0.017)	
$1 \mid 0.90 \leq$Firm federal R&D tax price$_t < 0.95$			0.21***
			(0.018)
$1 \mid$Firm federal R&D tax price$_t < 0.90$			0.24***
			(0.02)
Year FE	Yes	Yes	Yes
Firm FE	Yes	Yes	Yes
N	53,151	53,151	53,151
R^2	0.92	0.92	0.92

Note: This table shows the relationship between dummy variables for the federal tax price of R&D in a given region and log R&D. Observations are at the firm-year level. The tax price can take values between 0 and 1; the minimum value is 0.845 in our sample. If the tax price is close to 1, then the firm should not treat R&D differently than other expenditure. If tax price is below 1, R&D is less expensive than other expenditure because of the tax credit. In column 1, within-firm R&D is 18 percent higher when the firm's tax price is below 0.99 relative to the values close to one. Note, we use a 0.99 cutoff instead of 1 because there is a large mass of observations arbitrarily close to 1, which are unlikely to incentivize a large increase in R&D. Columns 2 and 3 show that in regions in which the tax price is below 1, R&D is monotonically increasing with the decrease in the tax price. Standard errors are clustered by firm. $*$, $**$, and $***$ denote significance at the 10%, 5%, and 1% levels.

Table A.8: First Stage IV Results Controlling for Past R&D

Dependent variable: Firm log R&D_t

	(1)	(2)	(3)	(4)	(5)	(6)
Firm log federal R&D tax price$_t$	-1.3***	-1.4***	-1.1***	-1.2***	-1.4***	-1.5***
	(0.083)	(0.084)	(0.089)	(0.091)	(0.18)	(0.19)
Firm log R&D_{t-1}	0.67***	0.65***	0.65***	0.64***		
	(0.011)	(0.011)	(0.013)	(0.014)		
Firm log R&$D^2{}_{t-1}$	0.023***	0.021***				
	(0.002)	(0.002)				
Firm log R&$D^3{}_{t-1}$	-0.003***	-0.003***				
	(0.0006)	(0.0006)				
Firm log R&$D^4{}_{t-1}$	0.0002***	0.0002***				
	(0.00006)	(0.00006)				
Firm log R&D_{t-2}			0.027*	0.027*		
			(0.014)	(0.014)		
Firm log R&D_{t-3}			0.0026	-0.014		
			(0.0099)	(0.01)		
Firm log sales$_t$					0.11***	0.098***
					(0.009)	(0.014)
Firm log sales$^2{}_t$					-0.008***	-0.006***
					(0.002)	(0.002)
Firm log sales$^3{}_t$					0.0003***	0.0003***
					(0.00009)	(0.00009)
Firm age$_t$					0.034***	0.034***
					(0.0062)	(0.0063)
Firm age$^2{}_t$					-0.0002***	-0.0002***
					(0.00004)	(0.00004)
Firm age$^3{}_t$					2.6e-07***	2.6e-07***
					(6.5e-08)	(6.5e-08)
Year FE		Yes		Yes		Yes
Firm FE	Yes	Yes	Yes	Yes	Yes	Yes
N	56,697	56,697	41,630	41,630	34,987	34,987
R^2	0.95	0.95	0.96	0.96	0.94	0.94

Note: This table shows that the federal tax price of R&D predicts R&D after controlling for past R&D or polynomials in age and sales. This tests the robustness of the first stage of the instrumental variables analysis (Table A.4). Sales are in millions of dollars. Age is defined according to the R&D tax credit legislation, where it is defined as the number of years from the first time the firm reported positive R&D. The requirement of positive age reduces the sample size for columns 5 and 6. The sample is an establishment-year panel of public firms. The federal R&D tax price is the log firm-level tax price of R&D, based on the federal tax credit, and following Hall (1993) and Bloom et al. (2013). Standard errors are clustered by firm. $*$, $**$, and $***$ denote significance at the 10%, 5%, and 1% levels.

Table A.9: Relationship Between State Tax Price of R&D and State New Firm Formation

Panel 1: Quarterly Workforce Indicator (LEHD) data

Dependent variable	Log 2-year employment growth		Change in 2-year-old firm total employment	
	(1)	(2)	(3)	(4)
State tax price of R&D	-0.74	0.33	-117	-6.5
	(0.59)	(0.36)	(7912)	(57677)
State FE	Yes	Yes	Yes	Yes
Year FE	Yes	No	Yes	No
N	449	449	448	447
R^2	0.44	0.43	0.11	0.11

Panel 2: Business Dynamics Statistics Data

Dependent variable	Log 2-year employment growth		Change in 2-year-old firm total employment	
	(1)	(2)	(3)	(4)
State tax price of R&D	-0.11	0.04	188	-583
	(0.37)	(0.08)	(1619)	(981)
State FE	Yes	Yes	Yes	Yes
Year FE	Yes	No	Yes	No
N	1530	1530	1529	1529
R^2	0.24	0.00	0.02	0.00

Note: This table shows estimates of the relationship between last year's state tax price of R&D (from Wilson 2009), and employment growth at new firms. Panel 1 uses data from the QWI, courtesy of Song Ma. Firms are limited to R&D-intensive (high tech) sectors. Panel 2 uses data from the BDS, where all firms are used because the data do not include industry information. Errors are clustered at the state level.

Panel 3				
Data source:	Quarterly Workforce Indicator (LEHD) data		Business Dynamics Statistics Data	
Dependent variable	Log 2-year employment growth	Change in 2-year- old firm total employment	Log 2-year employment growth	Change in 2-year-old firm total employment
	(1)	(2)	(3)	(4)
Federal R&D credit	4.4	-39912	-0.19	-377227
	(7.3)	(885697)	(0.16)	(274243)
N	16	15	30	37
R^2	0.03	0.00	0.05	0.05

Note: This panel shows estimates of the relationship between last year's federal tax price of R&D, and employment growth at new firms.