Entrepreneurial Spillovers from Corporate R&D

November 2018

Tania Babina and Sabrina T. Howell*

Abstract

This paper documents that corporate R&D investment increases employee departures to entrepreneurship. We use U.S. Census data, and instrument for R&D with its tax credit-induced cost. The ideas or skills that spill into startups seem to benefit from focused, high-powered incentives; for example, R&D-induced startups are much more likely to receive venture capital. The effect also seems to reflect ideas or skills that are poor complements to the firm's assets. As human capital is inalienable and portable, and startups are crucial to economic growth, R&D-induced labor reallocation to startups appears to be a novel channel of R&D spillovers.

*Columbia GSB, NYU Stern & NBER. Emails: tania.babina@gsb.columbia.edu; sabrina.howell@nyu.edu. The Kauffman Foundation funded Sabrina Howell's work on this project. For helpful comments, we are grateful to Manuel Adelino, Julian Atanassov, Pierre Azoulay, Nick Bloom, Shai Bernstein, Patrick Bolton, Alon Brav, Wes Cohen, Ryan Decker, Will Gornall, Bronwyn Hall, Wei Jiang, Bill Kerr, Saul Lach, Josh Lerner, Brian Lucking, Song Ma, Holger Mueller, Ramana Nanda, Stijn Van Nieuwerburgh, Paul Oyer, Tomek Piskorski, John Van Reenen, David Robinson, David Scharfstein, Philipp Schnabl, Antoinette Schoar, Xinxin Wang, Neng Wang, and seminar participants at Columbia, the Federal Reserve Board, the IMF, Duke, Drexel, Toulouse School of Economics, U.S. Census Bureau, NY Fed, Yale Junior Conference, NBER Corporate Finance Meeting, Stanford Financing of Innovation Summit, LBS Corporate Finance Conference, Stanford SIEPR, and Washington Summer Finance Conference. We also thank Chris Galvan and Shirley Liu for data assistance. The research in this paper was conducted while the authors were Special Sworn researchers of the U.S. Census Bureau at the Triangle Census Research Data Center. Research results and conclusions expressed are those of the authors and do not necessarily reflect the views of the U.S. Census Bureau. This paper has been screened to ensure that no confidential data are revealed.

1 Introduction

Corporate research and development (R&D) investment generates new knowledge, technology, and skills. Whether these intellectual assets are exploited within the firm or outside its boundary depends on their nature and on limits to contracting with employees (Aghion & Tirole 1994, Zingales 2000). R&D spillovers – in which R&D investment benefits other firms – are important for economic growth, but little is known about their transmission channel (Arrow 1962, Griliches 1992, Jones & Williams 1998, Klenow & Rodriguez-Clare 2005). Employees are an obvious channel for new knowledge and skills to leave the firm because their human capital is inalienable and portable. Yet there is little evidence about the effects of R&D on labor mobility across firms.[1]

In this paper, we ask whether R&D affects employee departures to entrepreneurship. To our knowledge, this is the first paper to establish a relationship between corporate R&D and human capital reallocation. We focus on departures to startups because they are conduits for new ideas and are important for economic growth.[2] Also, high-growth startup founders are often former employees of large incumbent firms (Gompers, Lerner & Scharfstein 2005, Klepper 2009). The effect of R&D on employee entrepreneurship is not obvious. In a frictionless environment, the firm would pursue all positive NPV projects that emerge from R&D and contract with employees ex-ante to prevent undesired departures to entrepreneurship. In the presence of contracting frictions, employees could take some R&D outputs to new firms. Alternatively, the effect could be negative if R&D leads to internal growth and better employment opportunities within the firm.

Testing the effect of R&D on employee entrepreneurship requires matching employers to employees and following the employees' subsequent career paths. We accomplish this with U.S. Census panel data between 1990 and 2008. We use the term "entrepreneurship" in a broad sense to mean the founding team of a new firm. Our main outcome variable is the share of an establishment's employees who depart and are among

[1] See Hall, Mairesse & Mohnen (2010).

[2] Entrepreneurs play a crucial role in prominent theoretical explanations for economic growth, including Schumpeter (1911), Lucas (1978), and Baumol (1990). Empirical literature has found that relative to incumbent firms, new firms have faster productivity and employment growth (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)).

the top five earners of a firm founded within three years. This captures founders and early employees - the group most likely to contribute new knowledge and crucial skills to the startup.[3] We find that a 100 percent within-firm increase in R&D leads to an 8.4 percent increase in the employee departure rate to entrepreneurship relative to the sample mean. Over the course of the sample, above- relative to below-median within-firm R&D changes yield 8,291 additional employee-founded startups, which is 7.7 percent of all employee-founded startups in the data. The model includes firm, state-year, and industry-year fixed effects, as well as time-varying firm characteristics. Despite fine controls, the estimate may be biased upwards if an unobserved new technological opportunity increases both parent R&D and and employee-founded startups. Alternatively, it may be biased downwards if the effect leads the parent to underinvest in R&D ex-ante.

To address these concerns, we instrument for R&D using changes in state and federal R&D tax credits, which affect the firm's user cost of R&D. We follow Bloom, Schankerman & Van Reenen (2013), but we provide new and exhaustive detail on the sources of within-firm variation for both instruments. The instruments satisfy the relevance condition and are likely to satisfy the exclusion restriction.[4] The instrumental variables (IV) effect of R&D on employee-founded startups is robust, offering evidence that the relationship is causal. It is about five times larger than the OLS estimate, which could reflect downward bias in the OLS result. Alternatively, the IV strategy estimates the marginal effect of R&D (the effect of an additional "last" dollar), while OLS gives the average effect (the effect of increasing the optimal amount of R&D by one dollar). The causal effect may be higher for the last dollar if it is spent on projects that are further from the firm's core focus or have less crucial outputs, and thus are more often rejected. It is also possible that adjustable R&D, the type sensitive to tax credit changes, has a larger causal effect. The true economic magnitude of the effect likely lies between the OLS and IV estimates.

This paper's main contribution is to establish the novel fact that firm-level R&D

[3]Similar variables are used in Kerr & Kerr (2017) and Azoulay, Jones, Kim & Miranda (2018), among others. The results are robust to a variety of alternative outcome variables, including the number of startups founded by recently departed employees.

[4]To satisfy the relevance condition, we present evidence from the literature that the elasticity of R&D spending to tax credits is at least one. To satisfy the exclusion restriction, we show empirically that there is no direct relationship between the tax credit and new firm creation, and present evidence from the legal literature that R&D tax credits are not in general useful to startups.

shocks leads to labor reallocation to startups. Regardless of the mechanism within the firm, our finding offers a dynamic source – changing internal R&D – for the location of the firm boundary around new growth options. It complements evidence in Robinson (2008), Rhodes-Kropf & Robinson (2008), Phillips & Zhdanov (2012), and Seru (2014) about the negative relationship between diversification and R&D productivity as well as the frictions that lead firms to locate innovative projects outside the firm, for example in strategic alliances.

While the data do not permit us to pin down a precise mechanism or channel for R&D-induced entrepreneurial departures, we explore support for theoretical implications of our finding. An idea or technology emerging from R&D investment will typically require substantial innovation investment to develop and commercialize. What may lead this process to occur in an employee-founded startup? Innovation is plagued by information, agency, and contracting frictions (Grossman & Hart 1986, Aghion & Tirole 1994). Contending with these frictions, the firm may opt not to pursue all good innovations. Some employee departures to entrepreneurship could be an unavoidable cost of R&D investment. One immediate implication is that R&D output over which the firm does not establish explicit property rights is most likely to yield employee-founded startups. Consistent with this, there is no effect of patents or patent citations on employee entrepreneurship.

Theories of the firm also yield two hypotheses. First, frictions are magnified when an idea is riskier, making high-risk, high-reward growth options more often best located outside the firm boundary (Gromb & Scharfstein 2002, Robinson 2008, Frésard, Hoberg & Phillips 2017). Many risky ventures benefit from the high-powered incentives that exist in small, focused firms financed with external capital markets. Consistent with this, we find that within the population of employee-founded startups, higher parent R&D is strongly associated with venture capital backing. Also, R&D-induced startups are more likely to be incorporated, more likely to be in high-tech sectors, have higher wages on average, and are more likely to 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.

Second, the effect may also reflect diversification costs, which lead the firm to sometimes reject R&D-generated growth options that are far from its core focus. There is less reason for physical assets that are not complementary to reside within the same firm (Williamson 1975, Hart & Moore 1990, Rhodes-Kropf & Robinson 2008). Indeed, we find that more parent R&D is negatively associated with the employee-founded startup being in same broad industry as the parent. We also show using supply chain relationships that R&D-induced employee-founded startups are more likely to draw inputs from a broader array of supplier industries. The types of ideas that spill into entrepreneurship seem to be those that benefit from focused, high-powered incentives and that are not especially complementary with the firm's existing activities.

We consider evidence for five alternative mechanisms. First, exposure to more R&D may increase an employee's entrepreneurial skills. This channel is likely at play, but some cross-sectional evidence is inconsistent with it being the main source of the effect. Second, employees may steal an idea that the firm values. If this were the primary driver, the effect should be smaller in states that strictly enforce non-compete covenants. Instead, those states exhibit a similar effect as states that weakly enforce non-competes. Third, the employee could cause the R&D increase or be hired as a result of it. We have strong evidence against both of these channels. Fourth, R&D may lead to restructuring, in which many employees depart the firm. Inconsistent with this, R&D does not lead to increases in other types of departures. Finally, the parent might fully internalize the R&D-induced startup's benefits if the startup is a wholly-owned spinout, in which case the effect would not be a spillover. We present evidence that parent firms do not appear to internalize the benefits of R&D-induced startups by investing in or acquiring them. In sum, while the data do not permit us to affirmatively identify a channel for our effect, the evidence best supports the two hypotheses grounded in the theory of the firm in which risky or diversifying growth options emerging from R&D more often end up in an employee's startup.

Entrepreneurial spillovers from corporate R&D may be costly to the parent firm; losing employees to startups is a "dark side" of R&D for the firm. While we do not assess the welfare effects of R&D-induced startups, our main 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. Acemoglu, Akcigit, Bloom & Kerr (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 somewhat different. By focusing on innovation inputs and labor reallocation, this paper extends the empirical literature on R&D spillovers, which includes Jaffe, Trajtenberg & Henderson (1993), Griffith, Harrison & Van Reenen (2006), Bloom et al. (2013), and Kerr & Kominers (2015).[5]

We offer corporate R&D as a new source for where ideas for high-growth startups come from, a topic of considerable recent interest (Aghion & Jaravel 2015, Babina 2017, Guzman & Stern 2017). There is existing evidence that successful entrepreneurs are often former employees of high-tech, large firms, and that employee-founded startups are related to agglomeration (Saxenian 1990, Gilson 1999, Bhide 2000, Klepper 2001, Gompers, Lerner & Scharfstein 2005).[6] To our knowledge, this paper is the first to document and quantify a causal effect of R&D investment on new firm creation by employees. More broadly, our paper is related to the literature on knowledge diffusion through labor mobility, including Almeida & Kogut (1999) and Herkenhoff, Lise, Menzio & Phillips (2018).

The paper proceeds as follows. We develop hypotheses in Section 2. Section 3 describes the data. Section 4 explains our reduced form and instrumental variables empirical approaches. The results are in Section 5. Section 6 discusses evidence for our hypothesized mechanisms as well as for alternatives. Section 7 concludes.

2 Hypothesis development

Theories of the firm offer predictions about how innovation interacts with firm boundaries, shedding light on whether a new technology or idea will stay inside the firm or move to a new, standalone firm. An idea emerging from R&D investment typically requires substantial innovation investment to develop and commercialize. When will it be the case that this development occurs in an employee-founded startup? We draw from two related

[5]The literature has typically assumed that potential recipients are close in technological or geographic space. Research at the individual level has focused on inventor networks, particularly in academia (Azoulay et al. 2010, Waldinger 2012).

[6]There is a rich management literature on employee-founded startups and spinoffs. For example, Klepper & Sleeper (2005) document within the laser industry that many new firms are founded by former employees of incumbent firms. Additional work includes Franco & Filson (2006), Klepper (2007), Hellmann (2007), Nanda & Sørensen (2010), Chatterji (2009), Sørensen (2007), Klepper & Thompson (2010), Campbell et al. (2012), Habib, Hege & Mella-Barral (2013), Agrawal, Cockburn, Galasso & Oettl (2014).

strands of theory. The first helps predict when a growth option should be located in a startup, and the second helps predict when the parent firm should determine that pursuing a new idea is NPV negative, even as it is NPV positive as a stand-alone firm founded by the employee. Both rely heavily on the prevalence of incomplete contracting. (As does an alternative story not considered in this section, in which incomplete contracting enables employees to steal ideas that the firm would prefer to retain in-house.)

The first perspective postulates that innovation investment is hard or impossible to contract on ex-ante, and innovation effort hard or impossible to verify ex-post (Grossman & Hart 1986 and Aghion & Tirole 1994). One implication is that some employee departures to entrepreneurship may be an unavoidable cost of R&D investment. Contracting and verification frictions also imply benefits to allocating residual rights of control to the party that performs innovation. As incentives to invest increase with control rights, integration is not always optimal. If the employee is responsible for the investment necessary to incubate an idea, effort may be optimal only in his own firm. Frésard, Hoberg & Phillips (2017) model these frictions to innovation explicitly in the context of vertical integration. They conclude that control rights should be allocated to stand-alone firms in especially R&D-intensive industries and when the innovation is as-yet unrealized; that is, when it requires more unverifiable effort. Acemoglu, Griffith, Aghion & Zilibotti (2010) also theorize that technology intensity should be associated with less vertical integration. This literature leads us to a possible dynamic relationship between R&D and employee-funded startups, summarized in Hypothesis 1.

Hypothesis 1: Corporate R&D has a positive effect on employee entrepreneurship.

Agency frictions are magnified when an idea is riskier, making high-risk, high-reward growth options more often best located outside the firm boundary. Gromb & Scharfstein (2002) model whether a new venture should be pursued within the established firm, "intrapreneurship," or outside the firm. They note that anecdotally, scientists and executives commonly leave large companies and launch their own ventures. 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. A different possible mechanism is elucidated in Robinson (2008). When embarking on risky projects, the firm cannot commit not to divert resources if the project fails. As a result, managers are unwilling to supply effort ex-ante. This makes it optimal to locate risky projects in a distinct legal entity outside the firm boundary. The firm can then contract with the new legal entity, committing not to "pick winners" ex-post. While our setting does not feature alliances as an outcome (in fact, we find that parent firms do not appear to benefit from R&D-induced startups), the underlying mechanism of inadequate effort provision helps explain why risky, diversifying ideas would leave the firm.[7]

In sum, in the presence of information asymmetry, agency problems and incomplete contracting, there are benefits to developing a risky new idea in a new venture rather than within the parent firm. This allocation of control rights implies that external capital markets, such as venture capital, will be better sources of financing than internal capital markets. This leads to our second hypothesis:

Hypothesis 2: R&D-induced employee-founded startups are more likely to be risky and potentially high-growth, because such ventures benefit from the incentive alignment inherent to small, focused firms.

The second theoretical strand concerns diversification. When a firm rejects a new idea that would diversify the firm's activities, employee-founded startups may be a byproduct. A starting point is the transaction cost-based theory of firm boundaries, premised on incomplete contracting. Williamson (1975) and Klein, Crawford & Alchian (1978) theorize that vertical integration reflects the importance of relationship-specific investments between transacting parties. Specific investments between separate firms create hold-up problems, or opportunities for one party to threaten to leave. Hart & Moore (1990) describe the firm as a set of property rights over physical assets, where residual rights of control (all non-contracted aspects of usage) reside with the owner of the asset. Only when physical assets are complementary will hold-up problems dictate integration within a single firm.

[7] Also see Lindsey (2008).

A more recent literature pushes forward this theory and links it to empirical facts. Rhodes-Kropf & Robinson (2008) demonstrate that when physical assets are more complementary, a merged firm will have greater surplus than two separate firms. A different mechanism is outlined in Seru (2014). He shows that information asymmetries between divisional managers and firm headquarters impede monitoring of divisional research and cause conglomerates to invest in less novel R&D. Also, the "picking winners" problem described above from Robinson (2008) is exacerbated for projects that would diversify the firm. The management literature has emphasized the complementarities mechanism as well. For example, Cassiman & Ueda (2006) and Hellmann (2007) argue that firms reject innovations that fit poorly with existing activities, which employees can then take outside the firm.[8]

Empirical work has found a negative correlation between firm performance and diversification (Lang & Stulz 1994, Schoar 2002).[9] 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). Efforts to explain diversification discounts have identified additional mechanisms, including the role of firm characteristics in the optimal degree of diversification (Campa & Kedia 2002, Maksimovic & Phillips 2002, Graham, Lemmon & Wolf 2002) and value-destroying behavior such as inefficient cross-subsidization (Scharfstein & Stein 2000, Rajan, Servaes & Zingales 2000). These explanations for equilibrium diversification are not inconsistent with a costly diversification mechanism explaining why a specific project might find itself optimally located outside the firm boundary.

In sum, if an R&D-generated idea is far from the firm's core focus and has weak complementarities, there is less reason for the new product to be integrated with the parent firm. Permitting the employee to take ownership and thereby residual rights of control may maximize investment incentives. In this case, we expect that R&D-induced startups would more often be in different broad industries from their parents. More generally, a permissive policy towards employee-founded startups could allow the firm to

[8] Using data from laser industry, Klepper & Sleeper (2005) focus on employee-founded startups that are also in the laser industry. They nonetheless find that these startups tend to target different customer segments than the parent.

[9] However, Whited (2001) and Villalonga (2004) argue that measurement error explains some of the discount evidence.

maintain the benefits of focusing on existing products and customers and could dynamically incentivize research employees to maximize effort. Thus, our third hypothesis is based on the idea of costly diversification.

Hypothesis 3: When a new idea or technology is further away from the firm's core focus, pursuing it will more often incur costs that exceed the benefits, leading R&D-induced employee-founded startups to more likely be in different markets than the parent and to draw inputs from a broader array of supplier industries.

3 Data

We use data from five sources: Compustat, Census LBD, Census LEHD, VentureXpert, and the NBER Patent Data Project. This section describes each source of data and explains the key variables we use in analysis. It also discusses potential concerns with the data.

3.1 Data Sources

Our measure of corporate innovation investment is R&D expenditure as reported in 10K filings and provided by Compustat. As R&D expenditure is only available for public firms, they form our universe of firms at hazard of being parents to employee-founded startups. We primarily use log R&D but show that the results are robust to using R&D divided by total assets. We restrict the sample to 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. Balance sheet and income statement data about the potential parents are also from Compustat.

We merge Compustat to the restricted-access U.S. Census Bureau's Longitudinal Business Database (LBD) using a Census-provided crosswalk. The LBD is a panel dataset that tracks all U.S. business establishments from 1978 to 2011 with paid employees, providing information on the number of employees and annual payroll. 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 (C- and S-corps rather than sole proprietorships or partnerships) comprise about 83 percent of the LBD.[10] For further details about the LBD, see Jarmin & Miranda (2002). We use the 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 employee departures to entrepreneurship is that we must observe employees and track them from firm to firm. We solve this with the Longitudinal Employer-Household Dynamics (LEHD) program at the U.S. Census Bureau, which provides quarterly firm-worker matched data. Abowd et al. (2009) describe the construction of the LEHD data in detail. 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 and over 96 percent of total wage and salary civilian jobs, so there is no problem with employee self-selection (BLS 1997, Abowd et al. 2009). About 10 percent of workers in year t are not in the LEHD in year $t + 3$, a similar attrition rate as the U.S. Current Population Survey.[11]

Coverage begins in 1990 for several states and increases over time, ending in 2008. We have access to 31 states, shown in Figure 1, in which we observe all employee-founded startups. These LEHD states are the location of 52 percent of the U.S. inventors, based on comprehensive USPTO patent data between 1990 and 2005. The LEHD data we use

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

[11] 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 percent drop out from the employment sample. Based on IPUMS-CPS data, available at https://cps.ipums.org/cps/.

also covers over 60 percent of U.S. employment, with representative firm age and industry composition. 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. To establish industry representativeness, we compare our data to data from the Bureau of Labor Statistics Current Employment Statistics Survey from 1990-2008.[12] We divide state-industry employment by total state employment across all states for each year, and then average across years. We conduct the same calculation for states out of our sample. The result is shown in Appendix Table A.1. For example, in the 1990-2001 period, the Manufacturing sector represents 15.4 percent of employment in our sample states, and 15.8 percent of employment in other states. In the 2002-08 period, the Professional and Business Services sector represents 12.3 percent of employment in our sample states, and 12.8 percent of employment in other states. 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.2. The share of employment for each industry is quite similar to the overall share of employment we observe.

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 ten employees, as they tend to have noisy reporting.[13] 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 venture capital-backed startups, we use Puri & Zarutskie (2012)'s link from ThomsonOne VentureXpert to the Census Business Register. We use patent data from the NBER Patent Data Project, which includes patent and citation variables through 2006. The NBER data include Compustat identifiers. We employ several annual patent-based variables at both the firm and industry level. These are the number of

[12]According to the BLS, employment data comes from a voluntary state level stratified sample of firms that is adjusted for population using monthly state unemployment insurance records.

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

patent classes a firm or industry patents in, the number of patents, the number of forward and backward citations, and the average, maximum, and median patent generality and originality. Generality is higher (closer to one than zero) when forward citations are in many classes, and originality is higher when backward citations are in many classes.

3.2 Identifying employee-founded startups

We are interested in measuring the effect of incumbent firm R&D on the departure rates of employees to entrepreneurship. Our final sample consists of an annual panel of public firm establishments in 31 states between 1990–2005. We measure departure rates at the establishment, as supposed at the firm-level, for two reasons. First, public firms often have operations in several industries and in several states. Aggregating to the firm-level would sacrifice this granular information on employees' industry and location. Establishment-level analysis permits including as controls industry-year and state-year fixed effects as well as establishment workforce characteristics and wages. 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.

To identify employee-founded startups, we begin by observing worker identities at public firm establishments in the first quarter of year t, and the quantity of R&D investment in year $t-1$. We denote an establishment e. Using longitudinally consistent individual identifiers in the LEHD, we follow the establishment e's employees one, two, and three years after year t. We follow startup creation from 1990 to 2008 because worker-level data are available over these years.

We proxy for an individual being on the founding team using the five highest earners at new firms. Our definition captures founders and the early employees who likely contribute crucial ideas and skills to the new firm. The measure is in line with prior research focusing on the executive team, including Gompers et al. (2005). Focusing on the highest earners not only captures workers with important human capital, but also likely captures the founders (Census data do not designate the founder(s) of a new firm). Kerr & Kerr (2017) show that a firm's top three initial earners usually include the firm's owners. As Azoulay et al. (2018) point out, 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.[14] 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 $t + 3$.[15] To arrive at our primary outcome variable – an establishment's rate of employee departures to new firms – we divide the number of founders by e's total number of employees in year t. For robustness, we show the effect on a range of other entrepreneurship measures (see Section 5.3).

There are four other future outcomes for the year t employees. First, they may remain at the firm. Second, they may be employed at a different firm that existed before year t (other incumbents). Third, they may be employed at an institution with unknown age (because some LEHD employers are non-profits, government entities, or non-employer firms not covered by the LBD, which is used to determine employer age). Finally, the employee may no longer be observed in the data, for example because he/she left the work force. We use these outcomes in robustness tests, for example to test whether R&D also leads to greater labor mobility to other incumbent firms.

3.3 Summary statistics

Table 1 Panels 1-3 show 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.[16] Our main dependent variable, employee entrepreneurship, is measured at the establishment-year level (Panel 2). This includes establishments of public firms with positive R&D and at least 10 employees between 1990 and 2005 (recall that the sample goes through 2008, but we allow three years to follow workers). On

[14]See https://www.irs.gov/uac/Wage-Compensation-for-S-Corporation-Officers.

[15]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 5.3.

[16]Since Census disclosure procedures prohibit disclosure of percentile value, we approximate 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.

average, 1.3 percent of an establishment's employees separate and are identified as entrepreneurs three years later. Similarly, Kerr, Kerr & Nanda (2015) find in the LBD/LEHD matched data that 1.7 percent of workers transition to entrepreneurship 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, with seventy percent being incorporated businesses. Two percent ever receive venture capital funding, which is much higher than estimates of the rate of venture capital backing among the whole population of new employer firms. Puri & Zarutskie (2012) find, also using Census data, that just 0.11 percent of new firms receive venture capital. Startups founded by recent employees of public firms with positive R&D are thus around eighteen times more likely to receive venture capital than the average firm.

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 Reduced form relationship

We estimate variants of Equation 1, where e denotes an establishment, f a firm, and t the year. As described above, the primary dependent variable is the percent of e_t's employees who are among the top five earners at startups as of $t+3$.

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

We employ firm fixed effects to control for time-invariant differences across firms. We expect omitted variables to be correlated within the firm, so we cluster standard errors by firm. Industry-year fixed effects (using SIC three-digit codes) control for changes in investment opportunities and subsume industry as well as year effects. We also use SIC four-digit codes in some specifications. State-year fixed effects control for regional shocks, which may affect investment opportunities at incumbents as well as entrepreneurship.

Time-varying establishment and firm controls address other concerns. First, we control for establishment size (number of employees) in case, for example, smaller establishments have more focused or autonomous cultures and thus lead to more employee entrepreneurship. Second, we control for the establishment's average wage in case R&D is associated with increases in wages. We also include the following firm-level controls, which might correlate with R&D and employee entrepreneurship: return on assets, sales growth, Tobin's Q, asset tangibility (measured as PPE investment divided by total assets), size (log total assets), cash holdings, age, and diversification (indicator for firm having establishments in multiple SIC three-digit industries).

4.2 Instrument for R&D

There are two major sources of endogeneity that may bias the ordinary least squares (OLS) estimates of Equation 1. On one hand, an unobserved demand shock or new technological opportunity not captured by the granular industry-year fixed effects may jointly engender parent R&D and employee entrepreneurship. This is a version of the Manski (1993) reflection problem and would bias the estimates upwards. On the other hand, the firm's inability to fully capture the benefits of R&D may reduce investment and bias the estimates downwards. This is one justification for the government subsidy of corporate R&D (Feldman & Kelley 2006, Howell 2017). We believe it is more likely that endogeneity biases the OLS result down. Two facts suggest that positive bias due to technology shocks is unlikely. First, adding industry-year fixed effects to specifications with firm fixed effects does not attenuate the estimates. Second, an opportunity shock in a given sector should lead to both more R&D and more startup formation in that sector. Instead, we find that the R&D-induced employee-founded startups and their parents tend to be in different sectors.

The ideal experiment would randomly allocate R&D to firms and observe whether firms assigned to more R&D have more employee entrepreneurship. This is infeasible, so we use the best available instrument: changes in the tax price of R&D following Bloom et al. (2013). While imperfect, we show that this IV strategy is well-suited to our context and is likely to satisfy the exclusion restriction.

4.2.1 Instrument motivation

We use two instruments: federal and state tax credit changes. Appendix Section A.1 contains exhaustive details about the credits, their calculation, and concerns with instrument validity. Here we summarize.

As with any instrument and accompanying reduced form estimation, 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 exclusion restriction, that the tax credits cannot directly cause employee entrepreneurship, is discussed in Section 4.2.4.) 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 at at least one, such that an extra dollar of credit stimulates roughly a dollar of additional R&D expenditure (or much more, in some studies). This evidence includes Hall (1993), McCutchen (1993), Mamuneas & Nadiri (1996), Hall & Van Reenen (2000), Billings et al. (2001), Bloom et al. (2002), Klassen et al. (2004), and Clausen (2009). In a particularly rigorous study, Rao (2016) finds that a 10 percent reduction in the user cost of R&D induced by the federal tax credit increases short-term (one-year) R&D spending intensity by about 20 percent. 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).[17]

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 3-4 percent in the longer term.

[17] 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, Rosell & Simcoe 2014, and Guceri & Liu 2017).

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 expenditure, 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.

R&D stimulated by tax credits is not simply a relabeling of existing related expenditure, but instead yields innovation and firm value creation. Dechezleprêtre, Einiö, Martin, Nguyen & Van Reenen (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. Further, Lucking (2018) uses cross-state variation to show that R&D tax credits increase employment growth, which does not come at the expense of employment growth in neighboring states. Lucking (2018) argues that the mechanism for the effect on employment growth is increased innovation.

4.2.2 Summaries of the tax credits

Changes in tax credits affect firm incentives to invest in R&D, because they change the firm-specific tax price of R&D (i.e., the 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. Here we briefly summarize the tax credits (see Appendix A.1 for details). The first instrument is the federal tax price of R&D, which we denote ρ^F_{ft}. The federal tax price has annual changes for most firms and is firm-specific for a number of reasons (Hall 1993). For example, it depends on firm age and past sales. It is calculated as a nonlinear function of these and other firm variables, so we can control for these variables directly in the IV. It is also worth noting that none of the variables on which the credit depends predict employee departures to entrepreneurship. We find substantial within-industry variation in the tax price of R&D, as well as the necessary variation within firm over time.

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. We use the state tax price of R&D in 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 the second object, $\theta_{f,s,t}$, we follow Bloom et al. (2013). $\theta_{f,s,t}$ is calculated using the share of the firm's patent inventors located in state s. The firm's state-level tax price is then $\rho^S_{f,t} = \sum_s \theta_{f,s,t}\rho^S_{s,t}$.

4.2.3 First stage estimation

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:

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

One potential concern is that R&D tax credits could have other effects in the state. As in the Equation 1, our IV estimation includes state-year and industry-year fixed effects. These will absorb any aggregate effects. We also continue to cluster standard errors by firm.

The results are in Table 3. The instruments are strong, yielding F-statistics of about 25, well above the rule-of-thumb cutoff of ten. 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).[18] As expected, the federal instrument is stronger than the state instrument. Bloom et al. (2013) use only firm and year fixed effects. This is equivalent to column 1. In column 2, we add firm time-varying controls, which reduce the magnitude of the effects somewhat but do not affect their statistical significance. 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. The results are also robust to using SIC four-digit industry fixed effects (column 6).[19]

[18] We expect that firms without taxable income would benefit less from R&D tax credits. Indeed, when we interact R&D with profitability (EBITDA/Assets), we find that while the independent effect of R&D remains positive and significant, the effect is significantly larger for more profitable firms. We do not report the estimates because of stringent limits on the number of estimates Census permitted us to disclose.

[19] 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.

4.2.4 Concerns with the instrument

There are five potential concerns with the instrument. The two more important ones are as follows (see Appendix A.1.3 for more details and the other three). First, the exclusion restriction is that tax credits cannot affect employee entrepreneurship except through the employer's R&D. We demonstrate in Appendix A.1.3 that there is no relation between the state tax credits and startup creation. More rigorously, Curtis & Decker (2018) show in a border-county differences-in-differences model that R&D tax credits have no effect on new firms. Further, the legal literature has argued that R&D tax credits are not useful to startups because they usually do not have taxable income (Bankman & Gilson 1999).[20] The fact that R&D tax credit changes do not affect a state's total number of startups is not inconsistent with our main result, which implies that overall corporate R&D may be an important source of high-growth startups. If R&D tax credits affect startup creation only through marginal corporate R&D, they are unlikely to yield a measurable effect on the state's aggregate number of startups.

The second concern is that changes in state-level R&D tax credits may lead firms to reallocate R&D (or misreport it such that it appears reallocated). Any such reallocation should reduce the power of the instrument. This leads us to expect that the federal instrument will have more power than the state instrument, which is indeed what we find. In sum, while imperfect, R&D tax credits offer the best available source of variation driving corporate R&D. Most importantly, they are plausibly unrelated to technological or demand shocks that could jointly give rise to parent R&D and employee entrepreneurship.

5 Main Results

This section first describes our main finding, which is causal evidence for Hypothesis 1. OLS and IV results are in Sections 5.1 and 5.2, respectively. Alternative measures of entrepreneurship are in Section 5.3. Alternative measures of R&D as well as parent firm heterogeneity are in Section 5.4. Alternative explanations for the main effect, such as firm restructuring, are addressed in Section 5.5. Reverse causation is examined in Section 5.6.

[20] The presence of carry-forwards 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.

5.1 Reduced form results

The main results from estimating Equation 1 are in Table 2. Our preferred specification in column 5 of panel 1 includes firm, industry-year, and state-year fixed effects. The coefficient of 0.109 implies that a 100 percent increase in parent firm R&D is associated with an 8.4 percent increase in the employee departure rate to entrepreneurship, relative to sample mean of 1.3 percent (a 100 percent increase in R&D corresponds to 1.1 standard deviation of R&D).[21] The result is robust to a wide array of alternative controls and fixed effects, shown across the eight models in Panels 1 and 2. For example, the result is robust to using SIC four-digit industry fixed effects (Panel 1 column 4 and Panel 2 column 1).

Our baseline set of firm-level controls are reported in Panel 1. 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.[22] We use alternative controls in Panel 2 columns 2 and 3. Column 2 employs establishment employee-level controls. Establishments with a higher share of white workers or foreign-born workers are associated with more employee entrepreneurship. Note that the results do not attenuate with wage controls, so the effect is not driven by an increase in employee wages. We include firm-level measures of patenting activity in Table 2 Panel 2 column 3. These are discussed in detail in Section 6.3.

5.2 Instrumented result

The results from the instrumented second stage are in Table 4. We repeat the specifications from Table 2. The coefficients in all models are statistically significant, and they are larger than the OLS results.[23] Our preferred specification, in column 5, is about five times the

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

[22] 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 (last quarter of year $t-1$), but establishment-level variables are measured when the employee snapshot is taken (first quarter of year t).

[23] 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 (Table 2 Panel 2 column 3). The firms responsible for the IV result are patenting in general, but changes in their number of

OLS estimate. The larger IV effect indicates that the subset of R&D expenditure affected by the tax credits leads to greater employee entrepreneurship than the average increase in R&D. This could reflect endogeneity that biases 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 possible explanations for such a phenomenon. One 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 or harder for the management to evaluate, perhaps because it is less predictable or farther from the firm's core focus.

The second possibility is a correlation between propensity to generate employee-founded startups and adjustable R&D. Adjustable R&D may be more general or inventive, and thus more often yield innovations best suited to development outside the firm. It is not obvious why adjustable R&D would be more inventive, but we cannot rule it out. More plausibly, adjustable R&D is less crucial to the firm. The loss of the innovation output to employee-founded startups would then be less costly, implying lower ex-ante incentives to prevent employee entrepreneurship. That is, suppose the firm expects R&D to lead to some employee-founded startups. When the loss of these employees and ideas is expected to be costlier, the firm should increase R&D less in response to the tax price shock.

If endogeneity biases the OLS result down, or if we capture the marginal effect of R&D better in the IV, then the IV better approximates the true effect of an independent increase in R&D. Conversely, if the IV isolates firms for which adjustable R&D is correlated with R&D-induced employee entrepreneurship, then the IV is biased upward. The true economic magnitude likely lies between the OLS and IV estimates.

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.

5.3 Alternative employee entrepreneurship measures

To demonstrate that our result is not sensitive to a particular construction of the outcome variable, we consider alternative measures of employee entrepreneurship in Table 5. 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 column 1 is the number of employee-founded startups from an establishment, normalized by employment at $t = 0$. We continue to observe a significant effect, indicating that team exits do not explain the main results. The coefficient implies that a 100 percent increase in R&D leads to a 5.8 percent increase in the number of employee-founded startups relative to the mean. Second, when only incorporated employee-founded startups are included, the effect is similar to the main effect (Table 5 Panel 1 column 2).

We continue to find a robust result using only the top three earners at the new firm rather than the top five (Table 5 Panel 1 column 3). The result is also robust to restricting employee-founders to those employed at the new firm in the first year it appears in the LBD with positive employment, for both the top five and top three earners (Panel 2 columns 1-2). In Panel 2 column 3, we consider only startups founded within one year (by year $t+1$ rather than $t + 3$). 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 in year two (Panel 3 column 1). When we consider one-year old startups in year $t + 3$, the effect is still positive, but becomes insignificant (not reported). Therefore, R&D-induced departures to entrepreneurship occur in the first two years after the investment in R&D.

We also replicate our main dependent variable using two instead of three years. We continue to find a significant effect (Panel 3 column 2). Finally, we use a "flow" measure of cumulative departures to entrepreneurship in Panel 3 column 3. Here entrepreneurs are defined as departed employees who are among the top five earners at a one-year-old employee-founded startup in year $t+1$, at a two-year-old employee-founded startup in year $t+2$, or at a three-year-old employee-founded startup in year $t+3$. The coefficient in this specification is also positive and significant at the .01 level. The results in this section are robust to using the IV estimator, though a few lose statistical significance at conventional levels. We do not report the IV results due to stringent limits on the number of estimates Census permitted us to disclose.

5.4 Parent heterogeneity and alternative R&D measures

We look for evidence that the effect is driven by establishments where R&D-generated ideas are likely to come from. Note that a new idea or technology need not leave the firm at its earliest stage. Instead, 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, however, we expect that R&D-generated ideas are more likely to be located in high-tech establishments. Industry is coded at the establishment-level, and there is substantial within-firm variation in establishment industries (among firms in our sample, the quasi-median is five three-digit SIC industries across establishments).[24] We interact R&D with a parent firm-level cross-sectional variable in Table 6.[25] Indeed, high-tech establishments drive our result, as the effect is 0.083 larger among high-tech establishments (column 1). There is no significant effect for non-high-tech establishments (the independent coefficient on R&D). This is consistent with Franco & Filson (2006)'s prediction that more technologically advanced firms are more likely to produce employee-founded startups. It provides comforting confirmation to our baseline results.

Patenting activity provides a second source of confirmation. General-purpose patents are used by 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 6 column 4). Also, recall that firms that patent in more classes tend to have higher employee entrepreneurship rates (Table 2 Panel 2 column 5). Thus, firms doing broader research have more employee-founded startups per dollar of R&D. Such research seems likely to yield ideas that are riskier or far from the firm's core focus.

Our results are robust to alternative measures of R&D, shown in Table 7. When

[24] We define an establishment as "high-tech" if its industry is in biotech, chemicals, software and business services, or high-tech manufacturing & R&D.

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

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.[26]

As robustness checks, Table 7 columns 3 and 4 use indicators for high and low changes in firm R&D. As expected, the effect is stronger when the independent variable is an indicator for the firm being the top 10 percentiles of R&D change (column 3). It 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.[27] All of the results in Table 7 are robust to using year-industry and year-state fixed effects as well.

R&D is observed at the firm level, but entrepreneurship is measured at the establishment level. We therefore implicitly assume that R&D is evenly distributed across establishments. The effect is driven by high-tech establishments, which is consistent with employees engaged in the R&D process driving the effect. Conceptually, as mentioned above, R&D spillovers need not come only from the establishments where R&D is actually performed; they might come from where R&D-generated projects or technologies are implemented, such as manufacturing plants, or where they are either rejected or pushed forward, such as headquarters.

[26]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.

[27]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.

5.5 R&D and employee turnover

R&D may lead to restructuring, in which many employees depart the firm. This could be an omitted variable creating correlation between R&D and employee entrepreneurship. Evidence against this hypothesis is in Appendix Table A.3. All of the results are robust to using the IV (we do not report them because of stringent limits on the number of estimates Census permitted us to disclose). First, column 1 shows that R&D in year t has no effect on the percent of employees who remain with the parent by year $t+3$. Column 2 finds that R&D has no effect on the percent of employees who move to another incumbent firm (firms that exist as of year $t-1$). Another concern is that parent R&D is correlated with worker mobility to or from uncovered state. R&D should then correlate with the fraction of workers who drop out of sample, but this is not the case. Columns 3 and 4 show that R&D has no effect on the percent of employees who drop out of the LEHD sample or move to organizations whose age is unknown. In addition to being statistically insignificant, the coefficients in all the regressions in this table are small relative to their means.

A second possible source of endogeneity is that R&D may lead the firm to hire new employees, who are inherently more likely to start their own ventures than the average worker. In this case, workers with relatively short tenures would drive the effect. In fact, we find that the effect of R&D on employee entrepreneurship is positive and significant among employees with above-median tenure, suggesting that workers hired specifically for the new R&D project do not drive the effect. (These regressions are unreported due to disclosure limitations.)

5.6 Reverse causation

If the effect of R&D on employee entrepreneurship is causal, employee entrepreneurship should not predict R&D. To test this, we project current-year R&D (in year t) on past employee entrepreneurship in Appendix Table A.4. 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 provides further evidence for causality of our main effect. In particular, it allays 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 the employee entrepreneurship occurs after the R&D.

6 Mechanisms

This section considers evidence for the latter two hypotheses stipulated in Section 2: R&D-induced employee-founded startups are more likely to be high-risk and potentially high-growth (Section 6.1), and they more often reflect costs to diversification (Section 6.2). We discuss how our patent results reflect incomplete contracting, which is crucial to both hypotheses, in Section 6.3. 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 2 concerns venture capital (VC) backing. VC-backed startups are widely known to be risky, associated with new-to-the-world ideas, and potentially high-growth.[28] We examine in Table 8 whether parent R&D is associated with certain startup characteristics. The dependent variable in Table 8 Panel 1 column 1 is one if the employee-founded startup receives VC. The coefficient on R&D is 0.007, significant at the .01 level. This implies that a 100 percent increase in R&D leads to a 35 percent increase in the chances that an employee-founded startup is VC-backed. 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 for intent to be a high-growth firm 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). We also expect high-risk, high-growth ventures emerging from R&D to be high tech. Indeed, they are more likely to be in a high-tech industry (Table 8 Panel 2 column

[28] Gornall & Strebulaev (2015) show that among U.S. public companies, those with VC are responsible for 44 percent of research and development expenditure, and Kaplan & Lerner (2010) show that over 60 percent of IPO issuers have VC backing.

2). Further, R&D induces employee-founded startups with higher wages than the average employee-founded startup, suggesting that they employ higher skill labor (Table 8 Panel 2 column 3). Finally, we consider the rate of exit, which we view as a proxy for risk. The vast majority of exits are probably firm failures, with a small minority being successful exits through 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 effect of 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. These startup-level results also serve to corroborate our main result. It would be less likely that R&D investment was the mechanism if we had we found that R&D were more likely to stimulate "Main Street"-type businesses such as restaurants or plumbing companies.

6.2 Costly diversification

The third hypothesis, a costly diversification mechanism, fits well with the IV interpretation where 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 driving 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 for the costly diversification hypothesis.

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 more parent R&D reduces the chances that the startup is in the same industry as its parent; a 100 percent increase in log R&D makes it 4.2 percent less likely that the employee-founded startup is in the same industry as its parent.

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 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. Second, in the 1990s, Michael Rosenfelt worked for the computer memory company Micron Electronics (now Micron Technology), where he helped to revitalize its PC business. He left in 1999 to found Powered Inc., an online education company, exploiting marketing innovations at Micron.[29] Micron Technology is in SIC 36 (Electronic and other Electrical Equipment), while Powered, Inc. would be in either SIC 73 (Business Services, the location of most Internet companies), or SIC 82 (Educational Services). Finally, David Friedberg and Siraj Khaliq left Google in 2006 to start WeatherBill (later The Climate Corporation), an agricultural insurance startup ultimately acquired by Monsanto.[30] Google's parent company Alphabet is in SIC 73 (Business Services), while WeatherBill would be in SIC 63 (Insurance Carriers). WeatherBill employed artificial intelligence insights from Google to better price insurance. In all three examples, an R&D-intensive parent spawned a new firm in a different two-digit SIC code sector, but the underlying idea was related to the parent's intellectual capital. These examples highlight how SIC assignments reflect the firm's market more than its technology. It seems likely that R&D-induced startups often employ innovation related to the parent's technology but apply it to a different market.

Seru (2014) proposes that information asymmetries between headquarters and divisional management helps to explain why conglomerates perform less productive R&D. We expect such information asymmetries to be more acute in large firms. Larger firms may face higher costs to diversification, more often rejecting a new innovation. We find that large firms, defined as having above-median total assets within a given year, drive the effect (Table 6 column 2). We do not find significant interactions between R&D

[29] Powered, Inc raised $8.5 million in VC and served clients such as Bloomberg.com Inc. It was acquired by Sprinklr, and internet company, and continues to exist as a standalone subsidiary. See here, here, and here.

[30] See here.

changes and measures of firm diversification, but this may reflect the fact that there is little variation in diversification measures within R&D-performing firms.

6.2.2 Supply chain relationships

To explore links between the startups and their parents, we consider supply chain relationships. We use the U.S. BEA annual input-output tables to create annual measures of supply chain closeness between the parent firm's industry and the startup's industry. The measures assign one party to be upstream and the other to be downstream. The first measure is "downstream closeness," which is the downstream industry's share of the upstream industry's product. The second measure is "upstream closeness," which is the upstream industry's share of what the downstream industry uses.[31] 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 effect of 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.[32] The effect of "upstream closeness" is negative, which means that the parent's product tends to make up a relatively smaller share of 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),

[31]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 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 available at https://www.bea.gov/data/industries/input-output-accounts-data.

[32]To the degree the spawn purchases from the parent, this does not imply that the parent benefits from the spawn. If both industries are competitive, the spawn can presumably buy from new supplier should just charge the market price for the input. If the spawn earns abnormal profits, there is no reason the parent should extract surplus from the new supplier.

implying that the R&D-induced startup's product tends to make up a somewhat larger share of parent's inputs.

Together, the results demonstrate a tie between R&D-induced startups and their parents. Relative to the average employee-founded startup's industry, the R&D-induced startup's industry buys a larger share of the parent industry's product and also supplies a larger share of the parent industry's inputs. Personal ties to the supplier firm may be part of what makes the employee's new human capital more valuable outside the firm. 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 is not optimal. This is consistent with the R&D-generated new venture being farther from the parent's core focus.

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 no effect of the number of patents or patent citations on employee entrepreneurship (Table 2 Panel 2 column 3). We also find no significant interaction between parent R&D and the number of patents or patent citations.

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).[33] 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

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

patented. Ellison, Glaeser & Kerr (2010), who also use this knowledge sharing measure and find weak effects, 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 about 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. Relatedly, Frésard, Hoberg & Phillips (2017) argue that vertical integration is more likely when an innovation is protected by patents. Similarly, Anton & Yao (1995) point out that the choice of an employee to take an innovation depends on there being no or weak property rights associated with it.

6.4 Alternative mechanisms

This section considers four alternative mechanisms beyond those based on the theory of the firm that we have thus far emphasized.

6.4.1 Entrepreneurial skills

R&D likely induces employee learning and skill development, which could make the employee more productive as an entrepreneur. This channel likely plays a role. However, three pieces of cross-sectional evidence suggest that it may not be the primary driver. First, in a human capital channel, we 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 6 column 2). Second, we might also expect that there is more opportunity for entrepreneurial learning at young firms. However, when we interact R&D with firm age, we do not find a larger effect in young firms, shown in Table 6 column 3.

Third, we expect that capital expenditure would have a similar effect on employee entrepreneurship if the channel were skills, because new capital investment is likely to create similar project management skills as R&D projects. Instead, Table 2 Panel 1 shows that there is no effect of total investment or PPE investment on employee entrepreneurship. In sum, while it is most likely that both human and intellectual capital explain why R&D

leads employees to start their own firms, the data are most consistent with the intellectual capital channel being dominant.

A related concern 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 2018). In Table 6 column 5 we interact R&D with an indicator for having had an IPO within the last three years. The interaction is insignificant, while the effect of R&D remains large and robust.

6.4.2 Idea stealing

With perfect information, the parent firm could appropriate all good innovations that emerge from R&D and contract with the employee ex-ante so that he will not depart to start his own firm. The mechanism of costly diversification – precisely because of contracting and information frictions – is almost certainly accompanied by some costs to the parent of employee entrepreneurship. Our finding may highlight a "dark side" to R&D investment from the firm's perspective.

We cannot calculate the magnitude of the costs to the parent firm of R&D-induced employee-founded startups. However, several pieces of evidence suggest that very costly stealing of ideas is not the main mechanism. If it were, we would expect the effect to be attenuated in states that enforce non-compete covenants. Non-competes restrict employees from working for a competing firm within the state for a specified period of time. It has been found that non-compete enforcement reduces local R&D spillovers (Belenzon & Schankerman 2013, Matray 2015), and reduces within-state inventor mobility (Marx, Singh & Fleming 2015). The main result persists in states that enforce non-competes, and there is no significant effect on an interaction between R&D and an indicator for being in a weak enforcement state. We do not report this result due to stringent limits on the number of estimates Census permitted us to disclose.

Second, if idea stealing is responsible for the effect, it should be attenuated when intellectual property is easier to protect, which makes it easier to contract on innovation effort. We do not find that the effect varies with a measure of industry patentability. Finally, there is a revealed preference argument. By virtue of observing the persistent phenomenon of R&D-induced employee entrepreneurship, the parent either chose not to develop the idea in house or chose not to take ex-ante 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 (see also Anton & Yao 1995).

6.4.3 Employee interaction with R&D change

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 the tax price of R&D, which the employee obviously 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

If the parent firm either partially owns or subsequently acquires R&D-induced startups, then the parent internalizes, or captures, some of the startup's private benefits. Full internalization (where the parent wholly owns the spinoff and captures all its benefits) would imply that the effect we observe is not a spillover. Some internalization may occur, but full internalization is unlikely. 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. This null result is unreported due to disclosure limitations. 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 do not usually invest in or acquire employee-founded startups.

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

7 Conclusion

The outcomes of R&D investment are uncertain, serendipitous, and difficult to contract on. This paper shows that some growth options generated by a firm's R&D process are reallocated from large incumbents to startups. 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 and the firm faces costs to diversification.

Much of the innovation literature focuses on innovation outputs, especially patents and patent citations. This paper takes a novel approach by examining a likely unintended consequence of R&D inputs. 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. Regardless of the costs to the parent firm, 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., Griffith, R., Aghion, P. & Zilibotti, F. (2010), 'Vertical integration and technology: theory and evidence', *Journal of the european economic Association* **8**(5), 989–1033.

Adelino, M., Ma, S. & Robinson, D. (2017), 'Firm age, investment opportunities, and job creation', *The Journal of Finance* .

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., Galasso, A. & Oettl, A. (2014), 'Why are some regions more innovative than others? the role of small firms in the presence of large labs', *Journal of Urban Economics* **81**, 149–165.

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.

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.

Anton, J. J. & Yao, D. A. (1995), 'Start-ups, spin-offs, and internal projects', *JL Econ. & Org.* **11**, 362.

Arrow, K. (1962), Economic welfare and the allocation of resources for invention, Technical report, National Bureau of Economic Research, Inc.

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

Azoulay, P., Zivin, J. G. & Wang, J. (2010), 'Superstar extinction', *Quarterly Journal of Economics* **25**.

Babina, T. (2017), 'Destructive creation at work: How financial distress spurs entrepreneurship'. Working Paper.

Babina, T., Ouimet, P. P. & Zarutskie, R. (2018), 'Going entrepreneurial? ipos and new firm creation'. Working Paper.

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.

Bhide, A. V. (2000), *The Origin and Evolution of New Businesses.*, Oxford University Press.

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.

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

Campa, J. M. & Kedia, S. (2002), 'Explaining the diversification discount', *The journal of finance* **57**(4), 1731–1762.

Campbell, B. A., Ganco, M., Franco, A. M. & Agarwal, R. (2012), 'Who leaves, where to, and why worry? employee mobility, entrepreneurship and effects on source firm performance', *Strategic Management Journal* **33**(1), 65–87.

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

Cassiman, B. & Ueda, M. (2006), 'Optimal project rejection and new firm start-ups', *Management Science* **52**(2), 262–275.

Chatterji, A. K. (2009), 'Spawned with a silver spoon? entrepreneurial performance and innovation in the medical device industry', *Strategic Management Journal* **30**(2), 185–206.

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.

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.

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.

Feldman, M. P. & Kelley, M. R. (2006), 'The ex ante assessment of knowledge spillovers: Government r&d policy, economic incentives and private firm behavior', *Research Policy* **35**(10), 1509–1521.

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

Franco, A. M. & Filson, D. (2006), 'Spin-outs: knowledge diffusion through employee mobility', *The RAND Journal of Economics* **37**(4), 841–860.

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

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.

Gilson, R. J. (1999), 'The legal infrastructure of high technology industrial districts: Silicon valley, route 128, and covenants not to compete', *NYUl Rev.* **74**, 575.

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* .

Graham, J. R., Lemmon, M. L. & Wolf, J. G. (2002), 'Does corporate diversification destroy value?', *The Journal of Finance* **57**(2), 695–720.

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', *The Scandinavian Journal of Economics* pp. S29–S47.

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. & Stern, S. (2017), *Nowcasting and Placecasting Entrepreneurial Quality and Performance*, Vol. 75, University of Chicago Press.

Habib, M. A., Hege, U. & Mella-Barral, P. (2013), 'Entrepreneurial spawning and firm characteristics', *Management Science* **59**(12), 2790–2804.

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

Hall, B. H., Mairesse, J. & Mohnen, P. (2010), Measuring the returns to r&d, *in* 'Handbook of the Economics of Innovation', Vol. 2, Elsevier, pp. 1033–1082.

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.

Hellmann, T. (2007), ‘When do employees become entrepreneurs?’, *Management science* **53**(6), 919–933.

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

Howell, S. T. (2017), ‘Financing innovation: evidence from r&d grants’, *American Economic Review* **107**(4), 1136–64.

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, 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.

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.

Klassen, K. J., Pittman, J. A., Reed, M. P. & Fortin, S. (2004), ‘A cross-national comparison of r&d expenditure decisions: Tax incentives and financial constraints’, *Contemporary Accounting Research* **21**(3), 639–680.

Klein, B., Crawford, R. G. & Alchian, A. A. (1978), ‘Vertical integration, appropriable rents, and the competitive contracting process’, *The journal of Law and Economics* **21**(2), 297–326.

Klenow, P. J. & Rodriguez-Clare, A. (2005), ‘Externalities and growth’, *Handbook of economic growth* **1**, 817–861.

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

Klepper, S. (2007), ‘Disagreements, spinoffs, and the evolution of detroit as the capital of the us automobile industry’, *Management Science* **53**(4), 616–631.

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

Klepper, S. & Sleeper, S. (2005), ‘Entry by spinoffs’, *Management science* **51**(8), 1291–1306.

Klepper, S. & Thompson, P. (2010), ‘Disagreements and intra-industry spinoffs’, *International Journal of Industrial Organization* **28**(5), 526–538.

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.

Lindsey, L. (2008), ‘Blurring firm boundaries: The role of venture capital in strategic alliances’, *The Journal of Finance* **63**(3), 1137–1168.

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.

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.

Manski, C. F. (1993), 'Identification of endogenous social effects: The reflection problem', *The Review of Economic Studies* **60**(3), 531–542.

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

Matray, A. (2015), 'The local innovation spillovers of listed firms', *Working Paper* .

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* .

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.

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

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.

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.

Rajan, R., Servaes, H. & Zingales, L. (2000), 'The cost of diversity: The diversification discount and inefficient investment', *The journal of Finance* **55**(1), 35–80.

Rao, N. (2016), 'Do tax credits stimulate r&d spending? the effect of the r&d tax credit in its first decade', *Journal of Public Economics* **140**, 1–12.

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.

Saxenian, A. (1990), 'Regional networks and the resurgence of silicon valley', *California management review* **33**(1), 89–112.

Scharfstein, D. S. & Stein, J. C. (2000), 'The dark side of internal capital markets: Divisional rent-seeking and inefficient investment', *The Journal of Finance* **55**(6), 2537–2564.

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.

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.

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.

Villalonga, B. (2004), 'Diversification discount or premium? new evidence from the business information tracking series', *The Journal of Finance* **59**(2), 479–506.

Waldinger, F. (2012), 'Peer effects in science: Evidence from the dismissal of scientists in nazi germany', *The Review of Economic Studies* **79**(2), 838–861.

Whited, T. M. (2001), 'Is it inefficient investment that causes the diversification discount?', *The Journal of Finance* **56**(5), 1667–1691.

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.

Zingales, L. (2000), 'In search of new foundations', *The journal of Finance* **55**(4), 1623–1653.

Table 1: Summary Statistics

Panel 1: Firm-year level variables

	Mean	Quasi-median	Standard deviation
Made corporate VC investments$_t$	0.038		
Had $\geq$ 1 patent$_{t-10,t}$	0.601		
Diversified$_{t-1}$	0.789		
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
Age$_{t-1}$	20.03	21.03	6.18
Total Assets$_{t-1}$ ('000s)	3,483	529	12,630
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)	0.613		
In high-tech industry	0.641		
Employee Entrepreneurship$_{t+3}$	1.31	0.82	2.43
# employee-founded startups$_{t+3}$	1.15	0.78	1.91
Stayers$_{t+3}$	47.77	52.30	25.98
Movers to old firms$_{t+3}$	26.29	22.51	18.10
Depart LEHD coverage$_{t+3}$	12.39	11.11	7.78
Movers to firms of unknown age$_{t+3}$	9.73	6.65	12.28
Average worker age$_t$(years)	40.08	40.27	4.76
Average employee tenure$_t$(years)	2.69	2.40	1.88
Share employees female$_t$	0.333	0.313	0.192
Share employees white$_t$	0.795	0.835	0.171
Share employees foreign$_t$	0.062	0.031	0.098
Number employees$_t$	329	122	1,698

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 within the interquartile range and a 1% weight on the remaining observations. R&D, assets, and wages are in real 2014 dollars.

Panel 3: Employee-founded startup level variables

	Mean	Quasi-median	Standard deviation
	(1)	(2)	(3)
Incorporated	0.698		
Same industry (SIC2) as parent	0.168		
Same state as parent	0.876		
High-tech industry	0.494		
Ever received VC	0.020		
Employee female	0.331		
Employee white	0.799		
Employee foreign	0.077		
Employee born in state	0.475		
Startup employment (in the first year)	11.83	5.41	29.85
Startup payroll (in the first year; '000s)	394	119	1,157
Startup age$_{t+3}$	1.59	1.99	1.01
Employee age$_t$	35.16	34.64	10.94
Employee education	13.66	14.36	2.49
Employee tenure (at parent firm; years)$_t$	2.07	1.58	2.25
Employee wages (at parent firm; '000s)$_t$	57.80	39.12	71.70
Employee wages (at employee-founded startup)$_{t+3}$	51.84	33.60	60.99

Note: Panel 3 shows summary statistics at the Employee-founded startup level. All variables are indicators and have 108,000 observations. Variables through "Employee born in state" are indicators, and the rest are continuous. "Employee" refers to individuals who left the parent firm to join the startup's founding team. Payroll and wages are in thousands of real 2014 dollars.

Table 2: Effect of R&D on Employee Entrepreneurship

Panel 1

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

	(1)	(2)	(3)	(4)	(5)
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)
Log employment$_t$		-0.217***	-0.181***	-0.174***	-0.179***
		(0.018)	(0.019)	(0.018)	(0.019)
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)
Sales growth$_{t-1}$		0.126	0.130	0.124	0.129
		(0.089)	(0.090)	(0.091)	(0.099)
EBITDA$_{t-1}$		0.131	0.127	0.155	-0.112
		(0.261)	(0.260)	(0.261)	(0.294)
Investment/Total assets$_{t-1}$		0.888	0.811	0.731	0.508
		(0.543)	(0.543)	(0.553)	(0.617)
Log Tobin's Q$_{t-1}$		0.022	0.032	0.027	0.044
		(0.066)	(0.067)	(0.067)	(0.077)
Log Total Assets$_{t-1}$		-0.011	-0.033	-0.054	-0.001
		(0.070)	(0.069)	(0.070)	(0.066)
PPE investment/Total assets$_{t-1}$		-0.177	-0.058	-0.050	-0.063
		(0.382)	(0.385)	(0.393)	(0.424)
Cash$_{t-1}$		-0.526*	-0.502	-0.506	-0.521
		(0.308)	(0.307)	(0.315)	(0.320)
Debt$_{t-1}$		-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)
Log R&D$_{t-1}$	0.102**	0.104**	0.101**
	(0.052)	(0.051)	(0.051)
Average employee age$_t$		-0.036***	
		(0.007)	
Share employees female$_t$		-0.084	
		(0.165)	
Share employees white$_t$		0.713***	
		(0.169)	
Share employees foreign$_t$		0.508**	
		(0.251)	
Average employee education$_t$		-0.055	
		(0.043)	
Average employee tenure$_t$		-0.023*	
		(0.013)	
Average employee experience$_t$		0.004	
		(0.017)	
Log patent classes$_{t-1}$			0.227*
			(0.120)
Log patents$_{t-1}$			-0.137
			(0.091)
Log forward citations$_{t-1}$			-0.006
			(0.022)
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 fraction of an establishment's workers as of first quarter of year zero who are entrepreneurs as of 1st quarter of year 3. An entrepreneur is defined as 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 3: First Stage IV Results

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

	(1)	(2)	(3)	(4)	(5)	(6)
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)
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 (Table 4). 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 7 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 diversified (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 4: Second Stage IV Results: Effect of R&D on Employee Entrepreneurship

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

	(1)	(2)	(3)	(4)	(5)	(6)
Instrumented 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 3. The dependent variable is the fraction of an establishment's workers as of first quarter of year 0 who are entrepreneurs as of 1st quarter of year 3. An entrepreneur is defined as 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 we are limited by Census in 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 diversified (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 5: Effect of R&D on Alternative Measures of Employee Entrepreneurship

Panel 1

Dependent variable:		Employee entrepreneurship$_{t+3}$	
	Number of employee-founded startups$_{t+3}$	Incorporated startups only	Top 3 earners (rather than top 5)
	(1)	(2)	(3)
Log R&D$_{t-1}$	0.067*	0.083**	0.099**
	(0.037)	(0.042)	(0.040)
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	0.154	0.183	0.154

Panel 2

Dependent variable:	Employee entrepreneurship$_{t+3}$		Employee entrepreneurship$_{t+1}$
	If employee present at startup founding	If employee present at startup founding and among top 3 earners (rather than top 5)	1-yr old startups only
	(1)	(2)	(3)
Log R&D$_{t-1}$	0.082*	0.081*	0.055**
	(0.048)	(0.040)	(0.025)
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	0.171	0.154	0.090

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 5.3. Controls are the same as in Table 2 Panel 1. Standard errors are clustered by firm. $*$, $**$, and $***$ denote significance at the 10%, 5%, and 1% levels.

Panel 3

Dependent variable:	Employee entrepreneurship$_{t+2}$ 1-yr old startups only	Employee entrepreneurship$_{t+2}$	Flow employee entrepreneurship$_{t+3}$
	(1)	(2)	(3)
Log R&D$_{t-1}$	0.057*	0.076*	0.89***
	(0.033)	(0.042)	(0.070)
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	0.097	0.131	0.209

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 5.3. Controls are the same as in Table 2 Panel 1. Standard errors are clustered by firm. $*$, $**$, and $***$ denote significance at the 10%, 5%, and 1% levels.

Table 6: Parent Variation in Effect of R&D on Employee Entrepreneurship

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

	(1)	(2)	(3)	(4)	(5)
Log R&D$_{t-1}$	0.048	0.016	0.035	0.099*	0.103**
	(0.057)	(0.062)	(0.066)	(0.052)	(0.052)
Log R&D$_{t-1}$·High Tech	0.083***				
	(0.029)				
Log R&D$_{t-1}$·Large$_{t-1}$		0.130**			
		(0.056)			
Log R&D$_{t-1}$·Old$_{t-1}$			0.098		
			(0.067)		
Log R&D$_{t-1}$·High patent generality$_{t-1}$				0.027*	
				(0.016)	
Log R&D$_{t-1}$·IPO$_{t-3,t-1}$					0.072
					(0.057)
Controls	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes
Firm FE	Yes	Yes	Yes	Yes	Yes
State FE	Yes	Yes	Yes	Yes	Yes
Industry (SIC3) FE	Yes	Yes	Yes	Yes	Yes
N	36,000	36,000	36,000	36,000	36,000
Adj. R^2	0.176	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 not. Large is 1 if the parent has above-median total assets (calculated at the firm-year level), and 0 if below-median. Old is 1 if the parent is of above-median age (calculated at the firm-year level), and 0 if below-median. High patent generality is 1 if the parent has above-median patent generality (calculated at the industry-year level), and 0 if below-median. IPO equals 1 if the firm went public within the past three years, and 0 otherwise. The dependent variable is the fraction of an establishment's workers as of first quarter of year zero who are entrepreneurs as of 1st quarter of year 3. An entrepreneur is defined as 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 2 Panel 1. Standard errors are clustered by firm. $*$, $**$, and $***$ denote significance at the 10%, 5%, and 1% levels.

Table 7: Effect of Alternative Measures of R&D on Employee Entrepreneurship

Dependent variable:	Employee Entrepreneurship$_{t+3}$	Number of employee-founded startups$_{t+3}$	Employee Entrepreneurship$_{t+3}$		
	(1)	(2)	(3)	(4)	(5)
Above median Δ R&D$_{t-1}$	0.089***	0.070***			
	(0.033)	(0.024)			
Top 10 pct Δ R&D$_{t-1}$			0.132**		
			(0.067)		
Bottom 10 pct Δ R&D$_{t-1}$				-0.105**	
				(0.053)	
R&D/Total Assets$_{t-1}$					1.020**
					(0.495)
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	0.175

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 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 fraction of an establishment's workers as of first quarter of year zero who are entrepreneurs as of 1st quarter of year 3. An entrepreneur is defined as 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. Controls are the same as in Table 2 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 Employee-founded Startup Characteristics

Panel 1: What predicts venture capital-backed employee-founded startups?

Dependent variable: Employee-founded startup ever received VC

	(1)
Log R&D$_{t-1}$	0.007***
	(0.001)
Employee age$_t$	0.001**
	(0.000)
Employee age$^2{}_t$	-0.000**
	(0.000)
Employee female	-0.013***
	(0.002)
Employee white	0.003**
	(0.001)
Employee foreign	-0.002
	(0.004)
Employee born in state	-0.007***
	(0.001)
Employee education	0.001***
	(0.000)
Employee experience$_t$	-0.000
	(0.001)
Employee tenure$_t$	-0.000
	(0.000)
Employee log earnings$_t$	0.008***
	(0.002)
Employee-founded startup age$_{t+3}$	0.007***
	(0.001)
Employee-founded startup initial employment	0.008***
	(0.002)
Continued...	
...Continued	
Establishment Log Employment$_t$	0.001
	(0.001)
Establishment average employee wage$_t$	0.012***
	(0.003)
Firm Age$_{t-1}$	-0.002***
	(0.001)
Firm Diversified$_{t-1}$	-0.003
	(0.006)
Firm Sales growth$_{t-1}$	0.004
	(0.006)
Firm EBITDA$_{t-1}$	-0.008
	(0.016)
Firm Investment/Total$_{t-1}$	-0.013
	(0.041)
Firm Log Tobin's Q$_{t-1}$	0.002
	(0.004)
Firm Log Total Assets$_{t-1}$	-0.006***
	(0.002)
Firm PPE Investment/Total Assets$_{t-1}$	-0.004
	(0.012)
Firm Cash$_{t-1}$	0.076***
	(0.015)
Firm Debt$_{t-1}$	0.009
	(0.007)
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 types of employee entrepreneurship. The sample is at the employee-founded startup level. Based on the main variable used in Table 2, 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...					
	is incorp.	in high-tech industry	log wages$_{t+3}$	exit$_{t+5}$	in same industry (SIC2) as parent	in same state as parent
	(1)	(2)	(3)	(4)	(5)	(6)
Log 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 (SIC3) FE	Yes	Yes	Yes	Yes	Yes	Yes
N	108,000	108,000	108,000	108000	108,000	108,000
Adj. R^2	0.080	0.102	0.318	0.083	0.206	0.053

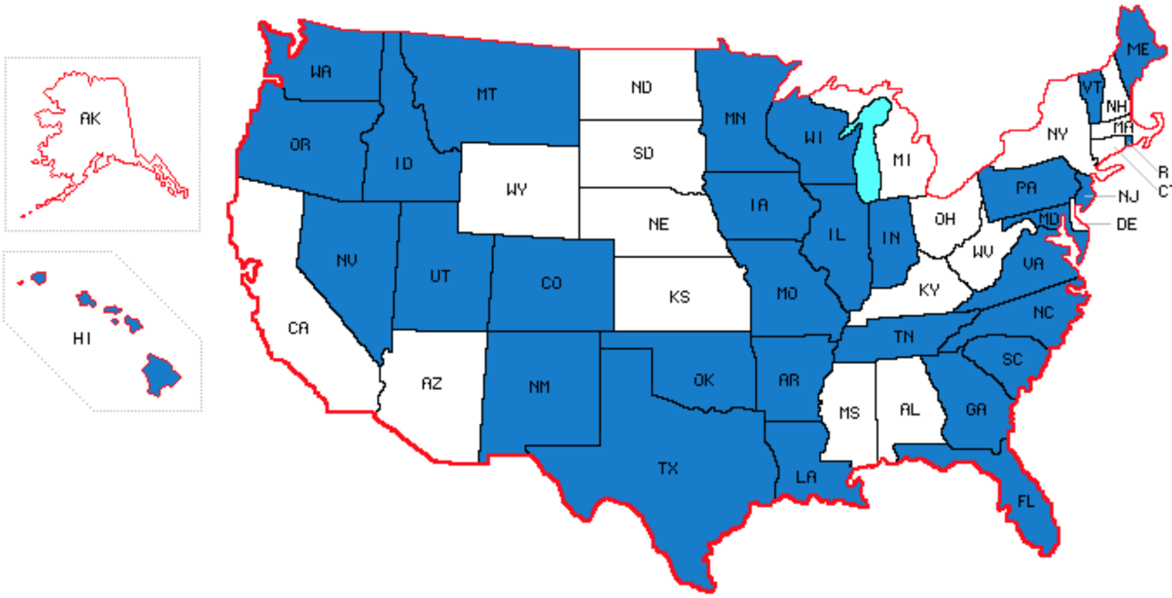
Note: This table shows the effect of R&D on types of employee entrepreneurship. The sample is at the employee-founded startup level. Based on the main variable used in Table 2, we identify whether the new firm associated with the departing employee has a given characteristic. The dependent variable in Panel 2 column 1 (2) (4) (5) (6) is 1 if the employee-founded startup is is an incorporated business (is in a high-tech industry) (the employee-founded startup exited (failed, though a small minority may be 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 3. An entrepreneur is defined as 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

	Parent upstream, employee-founded startup downstream		Parent downstream, employee-founded startup upstream	
Supply chain closeness measure:	Downstream closeness	Upstream closeness	Downstream closeness	Upstream closeness
Dependent variable:	Indicator for being in top 5% of closeness distribution			
	(1)	(2)	(3)	(4)
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.

Figure 1: Map of States with LEHD (Employee-founded Startup) Data



Note: This figure shows the 31 LEHD states that we have access to. We observe all employee-founded startups located in these states.

Appendix

(for online publication)

A.1 Instrumental variables calculation and discussion

A.1.1 The Federal R&D tax credit

The first instrument is the federal tax price of R&D, which we denote ${\rho_{ft}}^F$. 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 sample period for which we need to predict public firm R&D.[34] The general formula for the R&E tax credit is as follows, for tax year t and firm f:

$$R\&E\ Tax\ Credit\ Value_{tf} = 20\% \cdot [QRE_{tf} - Base_{tf}] + 20\% \cdot [Basic\ Research_{tf}] \quad (3)$$

The last element, basic research expenditures, must be paid to a qualified organization, which is either a research university or tax-exempt scientific organizations. The other, more complex type of research costs are qualified research expenditures (QRE). These must occur within the U.S., and have three categories: salaries and wages, supplies, and contract research. The law is quite specific about what counts and what does not count as QRE. For example, QRE must be technological in nature and relate to new or improved function, performance, reliability, or quality. Among other excluded types, research after commercial production of a component, survey research, and social science research do not count.[35]

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

[35] The complete legal text is here: https://www.law.cornell.edu/uscode/text/26/41.

The "base" amount is by far the most complicated element. It is constructed using the following equation:

$$Base_{tf} = Fixed\ Base\ \%_{tf} \cdot Sales_t$$

The complexity lies in the fixed base percentage, which varies by a firm's "startup" status. This term, which is used in the legislation and in Hall (1993), refers to the number of years since the firm's first instance of QRE. It is calculated as follows (firm index omitted for simplicity):

$$Fixed\ Base\ \% = \begin{cases} \max\left[\frac{\sum_{t=1984}^{1988} \frac{QRE_t}{Sales_t}}{5},\ 0.16\right] & if\ QRE_{1983} > 0\ \&\ Sales_{1983} > 0 \\ 0.03 & if\ QRE_{t-6} \in \{0, \emptyset\} \\ \frac{1}{6}\left[\frac{\sum_{t=-2}^{-1} \frac{QRE_t}{Sales_t}}{2}\right] & if\ QRE_{t-7} \in \{0, \emptyset\}\ \&\ QRE_{t-6} > 0 \\ \frac{1}{3}\left[\frac{\sum_{t=-2}^{-1} \frac{QRE_t}{Sales_t}}{2}\right] & if\ QRE_{t-8} \in \{0, \emptyset\}\ \&\ QRE_{t-7} > 0 \\ \frac{1}{2}\left[\frac{\sum_{t=-3}^{-1} \frac{QRE_t}{Sales_t}}{3}\right] & if\ QRE_{t-9} \in \{0, \emptyset\}\ \&\ QRE_{t-8} > 0 \\ \frac{2}{3}\left[\frac{\sum_{t=-4}^{-1} \frac{QRE_t}{Sales_t}}{4}\right] & if\ QRE_{t-10} \in \{0, \emptyset\}\ \&\ QRE_{t-9} > 0 \\ \frac{5}{6}\left[\frac{\sum_{t=-5}^{-1} \frac{QRE_t}{Sales_t}}{5}\right] & if\ QRE_{t-11} \in \{0, \emptyset\}\ \&\ QRE_{t-10} > 0 \\ \min\left[\frac{QRE_t}{Sales_t}\right]_{t-6}^{t-1} & if\ QRE_{t-x} \in \{0, \emptyset\}\ \&\ QRE_{t-x-1} > 0\ \forall\ \text{x} \geq 12 \end{cases}$$

In words, the first row is interpreted in the following way. For firms that had positive QRE and sales in 1983, the fixed base percentage is the maximum of 16% and the average of R&D intensity over the five years between 1984 and 1988. All the subsequent rows in the above equation pertain to what the law terms "startups." For example, for the first five taxable years after the first year in which a firm has positive QRE, the fixed base is 3%. In the 6th such year, it is one-sixth the average of the R&D intensity over the previous two years. The following rows are similarly calculated. Starting in the eleventh such year, firm may choose the percentage from any of the prior fifth through tenth years.

A few other details bear mention. The expense deduction for R&D is recaptured, reducing the effective credit rate from 20% to about 13.5%. Also, in the fiscal year 1995-6, the credit lapsed entirely. Additionally, when the credit value is larger that taxable profits, it can be carried forward for ten years. Finally, between 1990 and 1996, the only option was the R&E tax credit. Starting in 1996, firms could elect the alternative incremental credit (AIC), in lieu of the R&E tax credit. This has 3 tiers depending on R&D intensity (QRE relative to sales); if intensity is 1-1.5% (1.5-2%) (>2%), the AIC rate is 2.65% (3.2%) (3.75%), respectively. These rates have varied over time; they were lower in the late 1990s, and have increased in recent years.

The credit is firm-specific for a number of reasons. First, it depends on firm age, with annual changes for most firms. Second, the "base" amount of R&D is calculated using a firm's past R&D and current-year sales. Third, the base amount of the tax credit is the difference between realized R&D and the base. Fourth, there is a lower implicit value of the credit among tax exhausted firms because the value of the carry forward must be discounted. Finally, the lapse in 1995-96 generates additional within-firm variation, only for firms with R&D expenditures that year.

The R&E tax credit (denoted ERC_t) is in practice considerably more complicated to calculate than Equation 3, and follows Equation 7 in Hall (1992) and underlying equations not shown in her paper; these are available in Stata code on request. Calculating ERC_t begins with the tax credit rate (constant across firms), and multiplies by a categorical variable derived from QRE. This is then deducted from corporate tax liability. Then, a 3-year carry-back and a 15-year carry-forward are added in cases of no taxable income this year. Once this tax credit is arrived at, the tax price of R&D is calculated following Equation 6 in Hall (1992). This is:

$$\rho_{ft}^{F} = \rho_{t}^{R}\left[1 - T_t\left(1+r\right)^{-Jt}\tau\right] - \eta ERC_t \tag{4}$$

Here, ρ_t^R is an R&D deflator divided by a GDP deflator, or the "price" of R&D investment in the absence of taxes, T_t is an indicator for whether the firm has taxable income in the current year, J_t is the number of years until loss carry-forwards will be exhausted, τ_t is the corporate tax rate, and η_t is QRE. If ${\rho_{ft}}^F = 1$, then the firm should not treat R&D differently than other expenditure. If ${\rho_{ft}}^F < 1$, R&D is less expensive than other expenditure because of the tax credit.

In practice, we find substantial within-industry variation in ${\rho_{ft}}^F$, especially in manufacturing and services. The median tax price is well below 1 on average, so that R&D is cheaper than other spending. Within industries, the distributions have negative skew (i.e., a longer right tail). We also ensure that relevant current year variables, including R&D, do not have strong explanatory power over the tax price of R&D. Within firms, we find small positive correlations (all less than 0.1) between ${\rho_{ft}}^F$ and employment, assets, and R&D. In regressions, we verify substantial firm-level variation in the tax price of R&D. Firms in high tech areas such as pharmaceuticals and electronics, tend to have the most variation.

A.1.2 State R&D tax credits

State R&D tax credits have been generally modeled on the federal one. The first state R&D tax credit was implemented in 1982 by Minnesota; by the end of our sample period, forty states had some sort of R&D tax credit. The calculation of the base amount, and the 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}.[36] 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.[37] The firm's state-level tax price is then $\rho^S_{ft} = \sum_s \theta_{fst}\rho^S_{st}$.

A.1.3 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. 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.[38] The second is the Quarterly 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 counties being added over time.[39]

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 with BDS data are in Appendix Table A.5 Panel 2. We cluster errors by state. Regardless of the fixed effects or standard error assumptions, we find that the tax credits have no correlation with startup entry (Panel 1). Using the QWI sample, our dependent variable is either the logged new jobs created in new firms in the past two years, or the

[36]Specifically, it is roughly: $\frac{1-(tax\ credits+depr.\ allowances)}{1-tax\ rate}$.
[37]The data is from NBER patent data, available at
https://sites.google.com/site/patentdataproject/Home/downloads.
[38]This public version of the LBD is available at https://www.census.gov/ces/dataproducts/bds/data_firm.html.
[39]We used a transformed version of the data used in Adelino et al. (2017), courtesy of Song Ma.

change in the number of new jobs created in new firms in the past two years. We consider only R&D-intensive industries.[40] Again, regardless of whether we use year and/or state fixed effects, and regardless of the standard error assumptions, we find no effect of the tax price of R&D on these measures. This is in Appendix Table A.5 Panel 1.

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.5 Panel 3, again show no correlation.

More generally, the legal literature has argued that R&D tax credits are not useful to startups, as they have no or little taxable income against which to offset losses from failed R&D efforts (Bankman & Gilson 1999).[41] 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 reallocate R&D (or misreport it such that it appears reallocated). For studies 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 it 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

[40] NAICS codes 31-33, 51, and 54.

[41] 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."

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 firms choose to conduct less R&D in the years immediately preceding the tax credit change and more after in order to maximize the tax credit benefit. 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 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.

In sum, we believe that R&D tax credits offer the best available 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.

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.[42] 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.[43] Seventy-four percent of the unique startups matched to at least one investor or acquirer, yielding 20,478 unique startup-investor pairs.[44]

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.[45] 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.

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

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

[44] 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 (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.

[45] 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.

Others are asset managers, including Accel Partners, Softbank, 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, within our data we find no effect on employee entrepreneurship of the interaction between R&D and the parent having a corporate venture capital program. 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.

Table A.1: Sample Composition by Industry

Panel 1

1990 -2001

Industry	In Sample	Out 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

Industry	In Sample	Out 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.2: Industry Composition by Sample

Panel 1

1990 - 2001

Industry	In Sample	Out 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

Industry	In Sample	Out 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 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. 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.3: Effect of R&D on Non-entrepreneurial Employee Outcomes

Dependent variable:	Stayers$_{t+3}$	Movers to incumbent firms$_{t+3}$	Depart LEHD coverage$_{t+3}$	Movers to firms of unknown age$_{t+3}$
	(1)	(2)	(3)	(4)
Log R&D$_{t-1}$	-1.133	0.485	-0.004	0.506
	(0.715)	(0.608)	(0.133)	(0.452)
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	0.385	0.356	0.222	0.207

Note: This table shows the effect of R&D on alternative employee outcomes. The sample is an establishment-year panel of public firms. In column 1, the dependent variable is the fraction 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 column 2, the dependent variable is the fraction 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 3. In column 3, the dependent variable is the fraction 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 3 (note they may have moved to an uncovered state). In column 4, the dependent variable is the fraction of an establishment's workers in the 1st quarter of year zero who move to an organization whose age is unknown by the 1st quarter of year 3. Controls are the same as in Table 2 Panel 1. Standard errors are clustered by firm. $*$, $**$, and $***$ denote significance at the 10%, 5%, and 1% levels.

Table A.4: Reverse Causality Test (Effect of Employee Entrepreneurship on R&D)

Dependent variable: 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 current employee entrepreneurship does not predict 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 2 and 4. The one-year employee entrepreneurship$_{t-1}$ rate is the fraction of an establishment's workers as of first quarter of year $t-1$ who are entrepreneurs as of 1st 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 fraction of an establishment's workers as of first quarter of year $t-2$ who are entrepreneurs as of 1st quarter of year t. The three-year employee entrepreneurship$_{t-3}$ rate is the fraction of an establishment's workers as of first quarter of year $t-3$ who are entrepreneurs as of 1st quarter of year t. An entrepreneur is defined as 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 2 Panel 1. Standard errors are clustered by firm. $*$, $**$, and $***$ denote significance at the 10%, 5%, and 1% levels.

Table A.5: 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 as the data do not include industry information. Errors are clustered at the state *** indicates p-value<.01.

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. *** indicates p-value<.01.