Endogenous Growth with Employee Spinouts, Noncompetes, and Creative Destruction*

Nicolas Fernandez-Arias †
Princeton University
March 30, 2021
Click for most recent version

Abstract

I study the effect of noncompete agreements (NCAs) on aggregate productivity growth. I first construct a dataset matching venture capital funded startups to the previous employers of their founders. I find a statistically significant relationship between corporate R&D and subsequent employee startup formation in the same industry (within-industry spinouts). The relationship is economically significant, accounting for approximately 8.5% of startup employment in the data. Motivated by this finding, I develop a general equilibrium model extending a standard quality ladders model of endogenous growth to include R&D-induced within-industry spinouts and noncompete agreements. NCAs increase the incentive for R&D by incumbent firms while also stifling innovation by within-industry spinouts. According to the quantified model, reducing all barriers to the use of NCAs increases the annual growth rate by 0.21 percentage points and welfare by 3.24% in consumption-equivalent terms. It does so by improving the allocation of R&D labor, which in equilibrium is inefficiently overallocated to creative destruction due to business stealing and congestion externalities. Untargeted R&D subsidies reduce growth and welfare by exacerbating the misallocation of R&D. R&D subsidies targeted at own-product innovation can substitute for enforcement of NCAs without stifling spinout innovation. The growth-maximizing policy is a combination of large targeted R&D subsidies of 88% and a ban on the use of NCAs. When targeted R&D subsidies are lower than 62%, eliminating barriers to NCAs maximizes growth.

*I thank Esteban Rossi-Hansberg, Richard Rogerson, and Gianluca Violante for their continual guidance and support on this project. I also thank Ufuk Akcigit, Natalie Bachas, Gideon Bornstein, Oleg Itskhoki, Chad Jones, Ernest Liu, Adrien Matray, Ben Moll, Ulrich Müller, Ezra Oberfield, Alessandra Peter, Pascual Restrepo, Evan Starr, and Chris Tonetti. I also thank many participants at the Princeton Macro Workshop, the 2020 Economics Graduate Student Conference, the 2020 Young Economist Symposium, and the 2019 Wharton Innovation Doctoral Symposium. I am grateful for generous financial support from the Louis A. Simpson Center for the Study of Macroeconomics, the Bendheim Center for Finance, and the Julis-Rabinowitz Center for Public Policy and Finance. All errors are my own.

†Comments are appreciated and can be sent to nfernand@princeton.edu.
1 Introduction

The entry of new firms contributes significantly to innovation and productivity growth. Many innovative startups employ knowledge that their founders and employees gained while working at firms in the same industry. Such firms are known in the literature as within-industry spinouts (hereafter WSOs or spinouts) and there are many well-known examples. The most commonly cited example is the many spinouts of Fairchild Semiconductor, which include Intel and AMD. Fairchild itself is a spinout of Shockley Laboratories, another semiconductor firm. Moreover, former AMD employees founded NVIDIA, which competes with both Intel and AMD. Other examples can be found in many industries and time periods. Qualcomm, the leading designer and manufacturer of networking chips, was founded by former employees of Linkabit, a previously leading company in digital communications. In the software industry, Adobe was founded by former employees of the Xerox PARC research laboratory. The teleconferencing application Zoom was founded by a former vice president of engineering at Cisco Webex. SalesForce, a customer relationships management software company, was founded by a former executive of Oracle, one of its main competitors. In the pharmaceutical sector, Vertex Pharmaceuticals, which specializes in rational drug design, was founded by a scientist previously employed at Merck in the same capacity. The steel manufacturer Steel Dynamics – the third largest in the United States – was founded by former executives of Nucor, the largest steel producer. Examples can also be found outside of the high technology and manufacturing sectors. The management and business consulting firm Bain & Company was founded by a former vice president of Boston Consulting Group. In the financial sector, Bloomberg was founded by a former partner at Salomon Brothers who had set up internal computer systems.

To avoid competition by spinouts, incumbent firms may reduce investment in R&D and other forms of costly knowledge creation, as these activities expose employees to valuable knowledge and training which may facilitate the formation of a within-industry spinout. Alternatively, incumbents may attempt to prevent spinouts directly using a noncompete agreement (noncompete or NCA). When bound by an NCA, an employee is legally prohibited from founding a competing firm for a certain time after ceasing his or her current employment. Thus, NCAs allow employees to commit not to forming spinouts. While this of course inhibits innovation by spinouts, it also mitigates the disincentive to incumbent R&D. This economic tradeoff motivates several questions. What is the effect of NCAs on aggregate productivity growth? How does this depend on structural parameters that may be different in different locations, industries or time periods? Is it socially optimal to enforce NCAs?

In this paper, I take a step towards quantitatively answering these questions. I begin by examining the relationship between R&D spending and spinout formation. I first construct a dataset that matches parent firms to their spinouts. The dataset combines Compustat data on

---

1Vertex founder Joshua Boger’s goals for the company were to “become Merck, but better; design better drugs, faster; and become the 21st century biopharmaceutical company.” (see Werth, 2014)

2Other well-known examples in manufacturing can be found in the hard disk drive industry (Franco and Filson, 2006), the laser industry (Klepper and Sleeper, 2005), and the early automotive industry (Klepper, 2007).
publicly traded firms and private Venture Source data on venture capital-funded startups and their founders. Venture Source is the only dataset on startups with data on the previous employers of founders, executives and board members. As there are no common company identifiers, I use string matching techniques (e.g., regular expressions) to match Venture Source biographical data to Compustat firm name data. I define a startup as a within-industry spinout if its CEO, CTO, President, Chairman or Founder joined the startup in its first three years after being most recently employed at a firm in Compustat in the same 4-digit NAICS industry. Using this approach I match 2,475 founders of WSOs to their parent firms. Finally, I match this dataset to the NBER-USPTO database, which contains information on all US patents. In the resulting dataset, I find a statistically significant relationship between R&D spending at the firm level and subsequent within-industry spinout formation. This relationship holds when controlling for firm-level factors (including patents, a measure of the stock of knowledge) and using firm, firm age, industry-year and state-year fixed effects. It is also economically significant, accounting for around one tenth of employment in Venture Source startups.

I next develop a model which captures the relationship between R&D and spinout formation found in the data. It consists of a standard general equilibrium model of endogenous growth through creative destruction and own-product innovation augmented to include within-industry spinouts and NCAs. I assume that the aggregate supply of R&D labor is inelastic, focusing the analysis on its allocation between own-product innovation by incumbent firms and creative destruction by entrants. This assumption also allows for a closed form solution. To incorporate spinout formation, I assume that R&D employees for firms conducting own-product R&D stochastically gain the ability to form a within-industry spinout, replacing the incumbent.

To protect against the risk of being displaced by a spinout, incumbents can impose a noncompete on their R&D labor. Using an NCA, however, is not without cost. There is a direct cost, which represents the physical costs of enforcing the contract. In addition, in equilibrium, incumbents imposing an NCA must pay an endogenous wage premium in order to induce workers to voluntarily relinquish their right to profit from future spinout formation. In other words, when an NCA is not used, the incumbent implicitly compensates the employee through the expected value to the employee of future spinouts created. Of course, providing this implicit compensation is costly to the incumbent, as it implies a higher risk of losing profits. Incumbents choose whether to use an NCA in order to provide the equilibrium rate of compensation to R&D workers in the least costly way. The result of incumbents and employees optimizing is that NCAs are used exactly when they maximize the bilateral value of R&D employment.

In order for the model to generate a role for NCAs, therefore, there has to be a friction in the employment relationship that implies that, absent an NCA, a bilaterally suboptimal outcome may occur. Otherwise, by the logic in the previous paragraph, NCAs would be redundant. In particular, there is no role for NCAs if incumbents can acquire a bilaterally inefficient employee spinout before it is implemented and then shut it down (or implement it more efficiently inside the firm). This would prevent the bilaterally suboptimal outcome. While this “killer acquisition” strategy entails
a cost to the incumbent in the event an employee has an idea for a spinout, in equilibrium the employee would accept a lower wage \textit{ex ante} in anticipation of this buyout and, as a result, the incumbent would behave as though it had imposed a noncompete. To avoid this implication while keeping a simple model, I assume that there is no market in which spinouts can be sold to the incumbent firm that generated them. In reality, such a friction could be caused by asymmetric information concerning the quality of the idea (Chatterjee and Rossi-Hansberg, 2012), disagreements between the employee and the employer concerning the idea’s quality (Klepper, 2007), or limited commitment on the part of the employee (e.g., the employee cannot commit not to implement the idea even after selling it to his employer). In addition, antitrust law could prevent anticompetitive buyouts. Empirically, Babina and Howell, 2019 finds that only 2% of employee-founded firms in Venture Source are bought out by their former employers.\footnote{They consider employee-founded firms in Venture Source as identified in Gompers et al., 2005. They do not restrict attention to firms in the same industry as the parent firm.}

However, bilaterally inefficient spinouts need not be socially inefficient as innovation has positive externalities. Specifically, in the model innovation in one product increases the production and innovation efficiency of all other products in the economy. In this sense, NCAs contract the innovation possibilities frontier of the economy. On the other hand, NCAs can also increase growth due to their effect on the incentives for R&D by incumbents. In the model, NCAs increase the incentive for own-product innovation by incumbents, shifting R&D labor into potentially less efficient creative destruction by entering firms. For certain parameters, such firms have an incentive do R&D in equilibrium in spite of less productive technology. This occurs for two main reasons. First, the payoff per innovation through creative destruction is larger, as it replaces the monopoly of an existing incumbent. This is known as the business-stealing externality. In addition, they impose a congestion externality on each other as they attempt similar ideas without coordination.\footnote{I model this with a reduced form assumption following Acemoglu and Cao, 2015. The microfoundation is that when more firms attempt similar ideas, each firm’s chance to be the first successful firm declines rapidly while the expected time until the first firm succeeds declines marginally.} Overall, if R&D labor is sufficiently misallocated compared to the social value of within-industry spinout formation, enforcing NCAs can increase the equilibrium growth rate in spite of the fact that it contracts the innovation possibilities frontier.

To quantify the model, I calibrate it using the estimates from the empirical section, aggregate statistics, and innovative growth accounting decompositions from Garcia-Macia et al., 2019 and Klenow and Li, 2020. I also choose some parameters from the literature. I then use the model to study the effect on productivity growth and welfare of reducing the direct cost of using NCAs to zero, which I interpret as a relaxation of all legal barriers to their use. I find that growth rises by 0.21 percentage points and welfare rises 3.24% in consumption-equivalent terms. This results from the fact that, in the calibrated equilibrium, the business-stealing externality is strong and therefore R&D is substantially overallocated to creative destruction. Allowing the use of NCAs shifts R&D labor significantly to own-product innovation by incumbent firms. This corrects the equilibrium misallocation, increasing growth sufficiently to compensate for the reduced innovation by spinouts.
Next, I discuss how this result depends on parameters and, via the calibration, on the value of the targeted moments. I also discuss the sensitivity of the result to the parameters chosen externally. I show that the result is quantitatively robust to two plausible model extensions which attenuate it. Finally, I show an alternative calibration in which NCA enforcement reduces welfare and discuss the intuition for why the result is different in that case. In the alternative calibration, the employment share of young firms is lower while all other targets, including the growth share of young firms engaging in creative destruction, is held constant. This implies that each innovation by a new firm must be a larger improvement relative to the previous incumbent’s quality. In turn, this implies that a smaller share of the private return to creative destruction is business-stealing, mitigating the equilibrium misallocation of R&D spending. This weakens the growth increase from the reallocation of R&D labor such that growth is maximized by banning NCAs.

Finally, I consider some alternative policies that can substitute or complement the enforcement of NCAs. First, I consider the implications of R&D subsidies in this setting. I find that R&D subsidies can reduce growth by shifting R&D to entrants. In addition, they can induce incumbents to use NCAs when they otherwise would not have. Intuitively, R&D subsidies apply only to R&D wages, not to the cost of losing business to future spinouts. This makes R&D spending on own-product innovation by incumbents more expensive, shifting R&D labor to creative destruction by entrants in equilibrium. It also makes not using an NCA relatively more expensive, since the wage premium associated with using an NCA is subsidized. This induces the use of NCAs. Next, I consider targeted R&D subsidies; that is, subsidies that are only available to incumbent firm engaging in own-product innovation. I find that targeted R&D subsidies can significantly increase growth and improve welfare. However, just as with untargeted R&D subsidies, targeted R&D subsidies induce incumbents to use NCAs. Therefore, I finally consider the combination of targeted R&D subsidies and NCA enforcement policy. I find that the welfare-maximizing (growth-maximizing) policy is a large targeted R&D subsidy of 95% (88%) combined with a ban on the use of NCAs. However, if targeted R&D subsidies are below 77% (62%), reducing the barriers to the use of NCAs is optimal (growth-maximizing).

**Related literature** This paper adds to several strands of the literature. First, and most closely related, is a large literature examining the effect of noncompete contracts on worker mobility, entrepreneurship, incumbent firm investment and productivity growth. Due to the general lack of data on the use of NCAs, the empirical literature has typically exploited variation in the enforceability of noncompetes at the state-level. In some cases, this variation is argued to be exogenous, either due to legislative error (Marx et al., 2009, Marx et al., 2015) or due to unexpected court outcomes which then serve as judicial precedent (Jeffers, 2018). In addition, such papers typically use a diff-in-diff methodology exploiting a control industry that is unaffected by NCA enforcement, usually the legal industry. A survey of this work can be found in Bishara and Starr, 2016.

The results from the empirical work are consistent with the economic tradeoff of discussed above.

---

\[5\]I discuss some notable exceptions below.
Namely, higher enforcement of NCAs reduces worker mobility and entrepreneurship, but at the same time increases investment by incumbent firms. Regarding the effect on worker mobility, Marx et al., 2009 finds that inventor mobility declines in response to an exogenous increase in noncompete enforcement. Garmaise, 2011 finds that, in states where NCAs are more enforceable, managers are less mobile, have lower compensation, and invest less in their human capital. Balasubramanian et al., 2020 finds that a 2015 ban on technology worker NCAs in Hawaii increased job mobility and worker earnings. In an important qualification, Starr, 2018 finds that the negative wage effects associated with high NCA enforceability are driven by states in which workers can be asked to sign an NCA without “due consideration,” consisting of a wage increase or promotion.6 Regarding the negative effects on entrepreneurship, Stuart and Sorenson, 2003 finds less local entrepreneurship in response to local IPO in states that enforce NCAs. Samila and Sorenson, 2010 finds that an increase in VC funding supply increases entrepreneurship more in states without noncompete restrictions. Starr et al., 2018 finds that NCA enforceability is associated with reduced formation of WSOs, in particular preventing the formation of relatively low quality WSOs. On the other hand, higher NCA enforcement has been found to positively affects investment by incumbent firms. Conti, 2014 finds that higher enforceability of NCAs leads incumbent firms to pursue riskier R&D projects. Jeffers, 2018 finds that higher enforceability of NCAs increases investment by incumbent firms. In a related finding, Colombo et al., 2013 finds that easier access to finance leads to a reduction in incumbent firm knowledge investments.

While examining the effect of state-level differences in NCA enforcement provides valuable insights, some papers point to important qualifications of that approach. Marx et al., 2015 returns to the Michigan increase in NCA enforcement and finds evidence of significant brain drain of productive and collaborative inventors towards low-enforcement states in response. This suggests that differences in state-level outcomes by NCA enforceability – such as entrepreneurship and wages – could reflect reallocation across states of the human capital inputs to entrepreneurship, innovation, and productivity growth. Moreover, recent work has highlighted the importance of using individual-level data on the use of NCAs. Starr et al., 2019 constructs the first nationally representative dataset on the use of noncompetes at the employee level. They find that 20% of employees in the US are bound by an NCA, with substantial heterogeneity across occupations, industries and income groups. In particular, more than 35% of workers in computer, mathematical, architectural, or engineering occupations; more than 30% of workers in the professional, scientific and technical services and information industries; and more than 45% of workers earning more than $150,000 per year are bound by NCAs. This confirms anecdotal reports about the prevalence of NCAs in the United States. Puzzlingly, however, they find that the incidence of NCAs does not vary significantly with the cross sectional state-level variation in the enforceability of NCAs. In follow up work, Starr et al., 2020 finds that workers bound by NCAs have reduced job mobility, particularly towards competing firms, regardless of the enforceability of NCAs. The reduced mobility is driven

---

6The authors suggest that this would allows firms to use NCAs to increase their bargaining power by imposing an NCA after the worker has already rejected other job offers.
by a reduced job search rate in both regimes, while offers from poaching firms are only reduced in enforcing regimes. The authors interpret this as evidence of employee’s limited knowledge of the enforceability of NCA contracts. Overall, these results suggest that empirical work exploiting the cross section of NCA enforceability may not be capturing the effect of noncompetes.

The literature has also examined the effect of NCAs on growth using a structural approach (Franco and Mitchell, 2008, Shi, 2018, Baslandze, 2019). Closest to this paper is the recent work by Baslandze, 2019, which examines the effect of NCAs on spinout entry and growth in a similar framework. She also uses a general equilibrium model of endogenous growth with employee spinouts, using Compustat and NBER-USPTO patent data to discipline the analysis. There are several important differences. First, as she identifies spinouts using patent data, she focuses on spinouts whose founders patented at both the parent firm and the spinout. Second, her study focuses on all spinouts rather than within-industry spinouts. Thus she empirically considers all spinouts and, in her model, spinouts do not compete with the parent firm. Instead, the harm to the parent firm results from losing a valuable R&D manager who embodies the firm’s technological lead over its closest rival. Third, in her model the flow of R&D spending does not lead to the formation of spinouts; rather, it is the stock of knowledge at the firm that makes spinouts formed by the R&D manager more valuable, increasing the incentive of the R&D manager to form a spinout. Finally, she encodes the extent of NCA enforcement as a cost of spinout formation. By contrast, in my model NCAs only reduce the formation of within-industry spinouts and whether they are used is an equilibrium outcome responding to the bilateral inefficiency of spinout formation. To my knowledge, mine is the only general equilibrium endogenous growth model with this feature.7

This paper also relates to a literature examining the nature of spinouts, parent firms, and the effect of spinouts on their parents. This literature has found that within-industry spinouts leverage knowledge developed at parent firms (Klepper, 2002, Agarwal et al., 2002). Further, it finds that WSOs are both prevalent and particularly productive and fast-growing compared to non-WSO entrants (Muendler et al., 2012, Baslandze, 2019). It also finds that the entry of spinouts is associated with worse firm performance, due to increased competition as well as the loss of key employees (Wezel et al., 2006, Campbell et al., 2012, Agarwal et al., 2013). Regarding the spawning of spinouts, the literature has emphasized the role of strategic disagreements in the decision to spin out (Klepper and Sleeper, 2005, Klepper, 2007, Klepper and Thompson, 2010). Gompers et al., 2005 also studies Venture Source data, constructing a similar dataset to mine but for an earlier sample period. They focus on the permanent firm-characteristics that lead to the formation of spinouts, whereas I focus on the role of R&D spending choices by the parent firm. Most relatedly, in a contemporaneous paper Babina and Howell, 2019 examines the role of parent firm R&D spending in inducing the formation of employee spinouts. They find evidence of a causal relationship. My results are consistent with their findings. There are several differences in the data used and the approach taken. First, they use administrative data from the Census Bureau. While these data are

---

7As noted above, however, I do not explicitly model the source of inefficient creative destruction when there is no NCAs imposed.
superior in coverage to Venture Source and Compustat, there are two advantages to my approach. First, VC-funded startups are particularly important in generating productivity growth so it is important to confirm their findings for this type of firm. Second, the Census Bureau did not give them access to data from California or Massachusetts. It is important to confirm their findings for these two states as they have the most private R&D spending in aggregate and per capita terms. In terms of approach, they have an instrumental variables analysis based on instruments from Bloom et al., 2013. I choose not to use this instrument due to concerns about the exclusion restriction. Of course, this means I cannot directly infer causality from my empirical results. It is worth noting, however, that the IV estimate in Babina and Howell, 2019 is five times larger than the OLS estimate and that the authors conclude that their OLS estimates are likely more reflective of the average treatment effect of corporate R&D on spinout formation.

This paper relates further to a literature examining the interplay of invention, the market for ideas, and the boundaries of the firm (Arrow, 1962, Anton and Yao, 1994, Anton and Yao, 2002, Anton and Yao, 1995, Franco and Filson, 2006, Chatterjee and Rossi-Hansberg, 2012. Anton and Yao, 1995 considers how such frictions, in particular the risk of expropriation, can lead to employees forming spinouts rather than implementing their ideas inside the firm. Franco and Filson, 2006 develops a model in which employees learn from their employers and use this knowledge to form spinouts. Their setting shares with mine the the fact that employees implicitly pay for the knowledge they acquire while employed at the parent firm through lower equilibrium wages. However, in their settings spinouts do not directly compete with the parent firm, and further assumptions about the growth process imply that the equilibrium is Pareto efficient. By contrast, in my model, as is standard in the endogenous growth models discussed in the next paragraph, the equilibrium is generally not Pareto efficient due to both positive and negative externalities associated with innovation.

This paper also relates to the literature on endogenous growth and firm dynamics discussed in Romer, 1986, Romer, 1990, Grossman and Helpman, 1991, and Aghion and Howitt, 1992, Klette and Kortum, 2004, Lentz and Mortensen, 2008, Acemoglu and Cao, 2015, Acemoglu et al., 2018, and Akcigit and Kerr, 2018, among others. Most related is Acemoglu et al., 2018, which examines the role of the allocation of skilled labor in the determination of productivity growth. The main finding is that subsidies to incumbent R&D worsen the allocation of skilled labor by preventing the reallocation to higher productivity entering firms. Their framework is different along two important dimensions. First, they assume incumbents only engage in creative destruction, as in Klette and Kortum, 2004 and Lentz and Mortensen, 2008. This means that incumbents and entrants are both engaged in business-stealing, eliminating any possibility that business stealing incentives could create a misallocation of R&D spending. Second, they assume that incumbents stochastically enter an absorbing state of low innovation productivity, but still tie up skilled labor that could otherwise do R&D as they hire executives and managers. Together, these assumptions yield a model in which equilibrium overallocates skilled labor to incumbents that are unproductive at innovation. In my model, by contrast, incumbents are unique in their access to the own-product innovation technology
which, due to the lack of business-stealing incentives, is crowded out in equilibrium by relatively inefficient creative destruction. Therefore, there is significant positive growth impact from policies that make own-product innovation by incumbents relatively less expensive.

Finally, this paper relates to a large literature studying the private and social returns to R&D. The empirical literature includes seminal works such as Griliches, 1979, Griliches, 1992 and later studies in Bloom et al., 2013, Harrison et al., 2014, Mohnen and Hall, 2013, and Doraszelski and Jaumandreu, 2013. Jones and Williams, 1998 and Comin, 2004 examine this question using a structural approach. A survey can be found in Hall et al., 2010. This literature has typically found that the social returns to R&D are significantly higher than the private returns. This paper complements this work by emphasizing the heterogeneity of this breakdown across different types of R&D. It finds that R&D aimed at improving existing products has a much higher ratio of social to private return than R&D aimed at creative destruction.

The remainder of this paper is structured as follows. Section 2 presents the construction of the data and the empirical analysis. Section 3 develops a model of endogenous growth through own-product innovation and creative destruction which includes within-industry emplmoyee spinouts and noncompete agreements. Section 4 describes the calibration of the model. Section 5 presents policy counterfactuals. Section 6 concludes.

2 Empirics of R&D and spinout formation

In this section I document the empirical relationship between R&D spending and employee spinout formation which motivates the model of Section 3 and informs the calibration and quantitative analysis in Sections 4 and 5. I construct a new micro dataset by matching data from Venture Source, Compustat, and the NBER-USPTO patent database. Using this dataset, I analyze the microeconomic relationship between firm-level R&D spending and subsequent employee spinout formation using a regression analysis.

2.1 Data

2.1.1 Sources

VentureSource The data on startups comes from Venture Source (VS), a proprietary dataset containing information on venture capital (VC) firms and VC-funded startups. I use a subsample of the data for US-based startups founded between 1987 and 2009 which contain information on their founding year. The data cover 24,014 startups, 93,010 financing rounds, and 276,772 individual-firm pairs. For each financing round, the data contain information on valuation, amount raised, and status of the business at the time of the round, such as employment and revenue (with some missing

---

8Gompers et al., 2005 also match these two datasets using data through 1999. Since then, many startups have been added which were founded during the time period they study. This occurs because startup founding dates are added retroactively after the date of the first VC funding, which is typically several years after the startup is founded.

9At the start of this project, these data were owned by Dow Jones. They have since been sold to CB insights.
data). Most importantly for this analysis, the data contain employment biographies for each of the startup’s founders and key employees (C-level, high-ranking executives and managers) and board members. In this regard, Venture Source is unique among VC investment databases. Some summary information about the dataset is contained in Table A1. The dataset is described in detail in Kaplan et al., 2002 and Kaplan, 2016.

Compustat The data on R&D spending comes from Compustat, a comprehensive database of fundamental financial and market information on publicly traded companies. I consider a subsample consisting of all firms headquartered in the United States in operation at any point between 1984 and 2006, consisting of 20,534 firms. In addition to data on R&D spending, Compustat contains information on industrial classification as well time-varying firm accounting data and employment.

NBER-USPTO The data on patents comes from the NBER-USPTO database, which contains comprehensive information on all patents granted in the United States from 1976 to 2006 and comes linked to Compustat. I consider the subsample of patents assigned to US firms, consisting of 1,457,136 patents.

2.1.2 Construction of dataset

The first step is to define which startup employees should be classified as founders. The Venture Source data contain information on high level employees and board members. For the purposes of this study, however, only those employees whose human capital is crucial to the value proposition of the startup should be considered founders. I therefore restrict attention to employees who (1) have a job title related to the core operations of the firm (Founder, Chief, CEO, CTO or President) and (2) join the startup in its first three years since its founding date. When information on the individual's date of joining the startup is missing, I impute it as the founding date of the startup.10

Next, I extract information on previous employment using the employment biography data in VS.11 The results of this procedure are summarized in Table 1. The top twenty previous employers include several well-known technology firms such as IBM, Microsoft, Cisco Systems, Google, and Intel. Two consulting firms, Andersen Consulting and McKinsey & Company, also make the top twenty. Using string-matching techniques, I match these data to the firm name variable in Compustat.12

Having linked startups to their founders’ previous employers, the next step is to determine which of these startups compete with the parent firm. I accomplish this by using industrial classification as

---

10 Table A2 shows a breakdown of the most frequent job titles.
11 VS biographies are text fields in a structured format, allowing parsing by regular expressions.
12 This requires some effort to standardize names. The key difficulty is that company names are essentially unique proper nouns. Therefore, automated procedures based on natural language processing tend to fail, as those are designed to interpret regular speech. The only option, besides manual data cleaning, is to use regular expressions to trim “Inc.”, “Corp.” and variants thereof from each entry and attempt to match entries that way. I look for exact matches to previous employers in the VS data. For previous employers in VS that do not match with any names in Compustat, I check against the business segment names, available from the Compustat Segments database.
<table>
<thead>
<tr>
<th>Employer</th>
<th>Count</th>
<th>Employer</th>
<th>Count</th>
</tr>
</thead>
<tbody>
<tr>
<td>IBM</td>
<td>174</td>
<td>Stanford University</td>
<td>56</td>
</tr>
<tr>
<td>Microsoft</td>
<td>168</td>
<td>Lucent Technologies</td>
<td>46</td>
</tr>
<tr>
<td>Cisco Systems</td>
<td>122</td>
<td>AOL</td>
<td>44</td>
</tr>
<tr>
<td>Oracle</td>
<td>109</td>
<td>Motorola</td>
<td>42</td>
</tr>
<tr>
<td>Verizon</td>
<td>101</td>
<td>Andersen Consulting</td>
<td>40</td>
</tr>
<tr>
<td>Sun Microsystems</td>
<td>94</td>
<td>Nortel Networks</td>
<td>40</td>
</tr>
<tr>
<td>Google</td>
<td>79</td>
<td>MIT</td>
<td>40</td>
</tr>
<tr>
<td>AT&amp;T</td>
<td>76</td>
<td>McKinsey &amp; Company</td>
<td>39</td>
</tr>
<tr>
<td>Intel</td>
<td>70</td>
<td>Texas Instruments</td>
<td>38</td>
</tr>
<tr>
<td>Hewlett-Packard</td>
<td>64</td>
<td>Apple</td>
<td>37</td>
</tr>
</tbody>
</table>

Table 1: Top 20 previous employers for founders in VS data.

a proxy for product market.\textsuperscript{13} Compustat includes data on the self-reported NAICS industry of the parent firms. On the other hand, VS does not contain self-reported NAICS codes. However, it does document industry using a proprietary industry classification whose categories, for the most part, have the same names as corresponding NAICS 4 or 5 digit categories. For the minority of categories which do not have a direct counterpart in NAICS, I simply look for the 4 digit NAICS category that coincides most in terms of the meaning of the category. Using these industrial classifications, I identify a founder-startup as a WSO whenever the startup is in the same 4-digit NAICS category as its parent.

Following this procedure, approximately 20% of startup founders are matched to a public firm in Compustat. Of these, one third came from an employer in the same four digit NAICS industry as the startup.\textsuperscript{14} Figure 1 documents the joint distribution of parent industry and child industry, defined by 2-digit NAICS codes for easier visualization even though I use 4-digit NAICS industries to define WSOs. Because the raw joint distribution is too heavily concentrated in certain industries to be easily visualized in this way, I instead show the conditional distribution of child industry (column) conditional on parent industry (row). The dark diagonal line reflects the prevalence of WSOs. The diagonal entries are particularly intense in industries 32-33 (manufacturing) and 51 (information, which includes software publishing).\textsuperscript{15} This is consistent with the hypothesis that R&D is associated with within-industry spinout formation, as these industries conduct 85% of R&D during the sample period. Figure B1 shows the joint distribution the other way around, with the probability of parent industry conditional on child industry.\textsuperscript{16} The match is further documented in

\textsuperscript{13}VS does include a “Competition” variable listing competitors. However, this variable is only available for about 20% of startups.

\textsuperscript{14}The remainder were either most recently employed at a firm that is not publicly traded, and so is excluded from Compustat, or they are missed by the matching algorithm. This is likely due to (1) significantly different spellings or naming conventions between VS and Compustat, or (2) they worked at a subsidiary of a publicly traded firm, but the subsidiary is not named in the Compustat Segments database.

\textsuperscript{15}Appendix Table A11 contains the names corresponding to all two-digit NAICS codes.

\textsuperscript{16}In Figure 1, the dark vertical line at column 51 (Information) indicates that parent firms of all industries tend to spawn spinouts in that industry. Similar dark regions appear at columns 54 (Professional, Scientific and Technical Services), and 32 and 33 (Manufacturing). In Figure B1, the dark horizontal lines at 51 and to a lesser extend 32, 33,
2.2 Corporate R&D and spinout formation

2.2.1 Preliminaries

Figure 2 visualizes the relationship between corporate R&D and spinout formation in a binned scatterplot at the parent firm-year level. On the horizontal axis is average yearly real effective R&D spending by the parent firm in years $t - 2, t - 1, t$ and on the vertical axis is the yearly number of employees founding startups in years $t + 1, t + 2, t + 3$. Both variables are also demeaned at the parent firm-year level.

52 and 54 indicate that child firms of all industries tend to have founders from those industries.

17It is important to note that I do not replicate their findings. My algorithm finds fewer matches early on and more matches later in their sample period but always a smaller fraction of total founders. This may result from the fact that startups are only added to the data after they raise venture capital funding for the first time, but may have been founded years earlier.

18Real effective R&D spending is calculated from nominal R&D spending by first deflating using the R&D deflator and then further deflating by an index of cumulative productivity growth. This reflects the assumption that a given rate of R&D spending to GDP generates a given flow of new firms. This will be the assumption in the model developed...
Figure 2: Scatterplot of average yearly founder counts (restricted to same 4-digit NAICS industry) in \( t + 1, t + 2, t + 3 \) versus average yearly R&D spending in \( t, t - 1, t - 2 \). R&D spending is measured in millions of 2012 dollars (using the R&D deflator) and additionally deflated by productivity growth (also with 2012 as the base year). Finally, both R&D and founder counts are demeaned at the firm and industry-state-age-year levels.

The solid line shows the best fit of a straight line through the underlying scatterplot. This suggests that firm-level deviations of R&D are associated with future employee spinout formation.

### 2.2.2 Regressions

To properly assess the statistical significance of this relationship, as well as to control for other confounding factors, I conduct a regression analysis. Regarding confounding variables, the main concern is that time-varying firm-level variables that are not constant at the state-industry-age-year level could be associated with both R&D spending and spinout formation in both positive and negative ways. For example, investment opportunities specific to the firm’s technological niche within its state-industry-age-year would induce R&D spending and potentially within-industry spinout formation by its employees pursuing related ideas. On the other hand, poor management or other tribulations at the parent firm could reduce R&D spending and simultaneously induce employees to found competing firms. To control for such factors, I include include time-varying firm controls (employment, cumulative patents, assets, intangible assets, net income, sales, capital expenditures, and Tobin’s Q\(^{19}\)). Consistent with the demeaning in the scatter plot, I include fixed

---

\(^{19}\)In the regression in levels, I control instead for the firm market value; that is, Tobin’s Q times assets.
Table 2: The regressions above relate corporate R&D to the entrepreneurship decisions of employees. The dependent variable is average yearly number of founders joining startups in years $t + 1, t + 2, t + 3$. The independent variables are averages over $t, t - 1, t - 2$. Firm controls are employment, assets, intangible assets, investment, sales, net income, cumulative patents (weighted by all future citations), and firm market value (column 1) or Tobin’s Q (columns 2-4).

<table>
<thead>
<tr>
<th></th>
<th>(1) WSO4</th>
<th>(2) WSO4</th>
<th>(3) WSO4</th>
<th>(4) WSO4</th>
</tr>
</thead>
<tbody>
<tr>
<td>R&amp;D</td>
<td>0.24***</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>log(R&amp;D)</td>
<td>0.47***</td>
<td>1.84***</td>
<td>0.83***</td>
<td></td>
</tr>
<tr>
<td>Firm FE</td>
<td>Yes</td>
<td>Yes</td>
<td>No</td>
<td>Yes</td>
</tr>
<tr>
<td>Age FE</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
</tr>
<tr>
<td>Industry-Age FE</td>
<td>No</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
</tr>
<tr>
<td>Industry-Year FE</td>
<td>Yes</td>
<td>Yes</td>
<td>No</td>
<td>Yes</td>
</tr>
<tr>
<td>State-Year FE</td>
<td>Yes</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
</tr>
<tr>
<td>Controls</td>
<td>Yes</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Clustering</td>
<td>Industry, State Firm Firm Firm</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>R-squared (adj.)</td>
<td>0.61</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>R-squared (within, adj)</td>
<td>0.23</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>pseudo R-squared</td>
<td>0.34</td>
<td>0.47</td>
<td>0.35</td>
<td></td>
</tr>
<tr>
<td>Observations</td>
<td>56961</td>
<td>4254</td>
<td>7049</td>
<td>471</td>
</tr>
</tbody>
</table>

Standard errors in parentheses
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table 2 displays the results of the regression analysis. The first column is the regression in levels, given by

$$WSO_{it} = \beta R&D_{it} + \gamma X_{it} + \alpha_i + \xi_{j(i)t} + \sigma_{s(i)t} + \eta_{a(i,t)} + \epsilon_{it},$$

where $i$ and $t$ index firm and year, $j(i)$ and $s(i)$ are the industry and state of firm $i$, and $a(i,t)$ is the age of firm $i$ in year $t$. As in the scatterplot, the dependent variable $WSO_{it}$ is the (annualized) number of founders previously employed at firm $i$ joining startups in years $t + 1, t + 2, t + 3$. $R&D_{it}$ and the firm controls $X_{it}$ are calculated as moving averages over years $t, t - 1, t - 2$. For clarity, the units of $R&D_{it}$ are in billions. The parameters $\alpha_i, \xi_{j(i)t}, \sigma_{s(i)t}, \eta_{a(i,t)}$ represent the firm, industry-year, state-year and age fixed effects, respectively. Standard errors are multiway clustered at the state and four-digit industry levels.

---

20I do not consider state-industry-age-year fixed effects, as in the scatterplot, because this substantially reduces statistical power.
The result is statistically significant at the 1% level. The point estimate of 0.24 is quantitatively consistent with the slope of the line in the binned scatterplot in Figure 2. The robustness of the result to the inclusion of controls and firm, age, industry-year, and state-year fixed effects is encouraging. However, because the specification is in levels, industry-year, state-year and age fixed effects may be unable to control for unobservable time-varying confounding factors. Intuitively, in a levels specification each fixed effect is restricted to be constant for all in the corresponding fixed effect cell. However, time-varying industry- or state-level shocks that encourage both R&D spending and spinout formation are likely to affect firm outcomes more in absolute terms for larger firms.

To address this concern, the next three columns show the results of Poisson pseudo-Maximum Likelihood (PPML) regressions, with a specification of

$$\begin{align*}
WSO_{it} &= \exp \left\{ \beta \log R&D_{it} + \gamma x_{it} + \alpha_i + \xi_{j(i)t} + \sigma_{s(i)t} + \eta_{a(i)t} \right\} + \epsilon_{it},
\end{align*}
$$

where $x_{it}$ denotes the logarithm of the firm-level controls. This specification can be thought of as a log-linear regression which can handle zeros in the dependent variable. It is also the specification in the model, as the rate of employee departures is assumed to depend on R&D spending following a Poisson process. In this specification, the industry-year, state-year and age fixed effects have a proportional effect on the formation of WSOs. This should allow them to better proxy for such unobservable effects which are correlated with R&D and spinout formation and hence can bias the results of the levels regression. I perform three regressions with this specification: one with controls, but no fixed effects (column 2); one with fixed effects but no controls (column 3); and one with both controls and fixed effects (column 4). In all cases, in order to have sufficient clusters, I cluster standard errors by firm rather than by industry and state as in the levels regression.

In all three regressions the coefficient on the logarithm of R&D spending is statistically significant at the 1% level. The coefficient on R&D spending corresponds to an elasticity of spinout formation to R&D spending. In column 2, which has no controls but includes firm, industry-age, industry-year, and state-year fixed effects, the the estimate corresponds to an elasticity of 0.47. This conforms to the elasticities estimated in the micro literature for other outputs of R&D such as patents. In column 3, which uses no fixed effects but includes the full set of firm controls, the estimate corresponds to an elasticity of 1.84. Finally, the estimate in column 3 uses both controls and fixed effects, exchanging industry-age fixed effects for age fixed effects in order to lose fewer observations. The estimate in this case corresponds to an elasticity of 0.83, somewhat less precisely estimated than the previous two due to the smaller sample size.

Additional results in Table A4 and Table A5 show that having including firm age, industry-year, and state-year fixed effects in the PPML regression did not reduce the coefficient estimate on the logarithm of R&D. In fact, when controls are not included, going from only firm fixed effects to the full set of fixed effects increases the estimate slightly increases from 0.4 to 0.47. When controls are

---

21 Even if Poisson is not the correct specification, the PPML specification consistently estimates the conditional mean of the actual distribution.

22 See Akcigit and Kerr, 2018 for a discussion.
included, the estimate increases dramatically from 0.5 to 2.17, although this may be due to the fact that many observations are dropped. These results suggest that correlated industry-year, state-year and firm age shocks to R&D and spinout formation are not responsible for the correlation between R&D and WSO formation. However, without exogenous variation in R&D, I cannot definitively rule out the possibility that the problematic shocks operate at a more granular level. For example, they could operate at the level of specific technological niches operated by the parent firm.

To get a sense of economic magnitude, I use the estimate in column 1 to construct an estimate of the fraction of WSO founders in the data that can be accounted for by parent firm R&D spending. In each year $t$, I compute $\hat{WSO}_t$, the expected number of within-industry spinout founders induced by R&D per year over the years $t+1, t+2, t+2$, by multiplying average aggregate real effective R&D spending in years $t, t-1, t-2$ by the coefficient estimates from column 1. I then aggregate across years in the sample. Based on this measure, the regression coefficient is economically significant: it can account for 85% of the WSO founders observed in the data. \(^{23}\) Further, Appendix tables A6, A7, and A8 document that startups with a higher fraction of WSO founders tend to have roughly 35% higher employee count, 45% higher revenue, and 36% higher valuation on a per-founder basis. This relationship holds after controlling for industry, state, time, cohort, and / or age factors,\(^{24}\) and is statistically significant and robust across specifications.\(^{25}\) Combining these last two observations with the fact that about 7% of startup founders are at WSOs suggests that R&D-induced WSOs account for about 8.5% of the level of employment, revenue and valuation of startups in the dataset.

Based on the results above, I conclude that the data are consistent with an economically significant effect of corporate R&D spending on within-industry spinout formation. While the absence of exogenous variation in R&D spending prevents any definitive claims of causality, the results above justify the development and analysis of a structural model which exhibits an effect of R&D on within-industry employee spinout formation.\(^{26}\)

\(^{23}\)Figure B2 shows a similar accounting exercise using data at the industry-year level, focusing on NAICS industries 3 (manufacturing) and 5 (information), which are responsible for the vast majority of private sector R&D spending. The pattern is clearly different by industry, with a very good fit for manufacturing industries but underestimated spinouts in information industries. A framework that is able to account for variation in this mechanism across industries would therefore be useful in future work.

\(^{24}\)I only control for two of the three (time, cohort, age) factors, to avoid the well-known multicollinearity problem.

\(^{25}\)Using the same methodology, I also find in tables A9 and A10 that startups with a higher fraction of founders from the same industry are significantly less likely to go out of business and more likely to be acquired or IPO. While I do not use this fact explicitly in my quantification of the model, it fills in the overall picture in a way consistent with the other findings.

\(^{26}\)I do not pursue an instrumental variables identification strategy. The natural choice for an instrumental variable would be the one described in Bloom et al., 2013. Its variation is largely driven by R&D tax incentives at the state-level, which may fail to satisfy the exclusion restriction in this case because WSOs typically locate in the same state as the parent firm and may also be sensitive to R&D tax incentives. The instrument is considered in Babina and Howell, 2019. The authors argue that R&D tax incentives play little role in startup formation as they cannot be claimed if the startup has no profits. They can be carried forward twenty years, although of course if the profits never materialize they cannot be claimed. In any case, the IV estimate is five times larger than the OLS estimate and the latter is the authors’ preferred estimate of the average treatment effect of R&D spending on spinout formation.
3 Model

Motivated by the empirical findings above, in this section I develop a model of endogenous growth with R&D-induced spinouts and noncompete agreements. It extends a standard quality ladders model of endogenous growth through own-product innovation and creative destruction, drawing mostly from Grossman and Helpman, 1991, Acemoglu and Cao, 2015, and Akcigit and Kerr, 2018. The new features are that own-product R&D spending induces WSO formation and, relatedly, that incumbents can use NCAs to prevent this from happening.

3.1 Representative household

I model a continuous time economy, starting at $t = 0$. The representative household has CRRA preferences over consumption streams of the final good $\{C(t)\}_{t \geq 0}$, given by

$$U(\{C(t)\}_{t \geq 0}) = \int_0^\infty e^{-\rho t} C(t)^{1-\theta} - \frac{1}{1-\theta} dt. \quad (3)$$

In each period $t \geq 0$, the household is endowed with $\bar{L}_{RD} \in (0, 1)$ units of R&D labor as well as $1 - \bar{L}_{RD}$ units of production labor which is used in the production of intermediate and final goods. The labor resource constraints are

$$L_{RD} \leq \bar{L}_{RD}, \quad (4)$$

$$L_P \leq 1 - \bar{L}_{RD}. \quad (5)$$

The household takes as given profits it receives from the ownership of all firms in the economy. The household also has access to an instantaneous risk-free bond that exists in zero net supply. Note that the above specification implies that the supply of R&D labor is inelastic. This enables an analytical solution to the model. It also implies that productivity growth is determined by the allocation of R&D labor rather than the amount. In reality, of course, the aggregate R&D labor allocation in the economy can vary. The present model can be viewed as an approximation that is more accurate in the short to medium term, when the total amount of R&D labor is constrained by the slow process of human capital accumulation. Alternatively, it can be interpreted as a stylized model which highlights the effects of R&D incentives on the allocation of R&D labor to different uses rather than the total amount of R&D labor supplied.\(^{27}\)

\(^{27}\)It would be straightforward to extend the model to a case with an endogenous aggregate supply of R&D labor, at the cost of no longer having closed form expressions for the resulting equilibrium allocation. Assume that the household has an underlying stock of labor $\bar{L}$, normalized to $\bar{L} = 1$, of which it deploys $\bar{L}_{RD}$ to the production of R&D human capital with concave production function $f()$. The remainder can be used for production. The R&D labor resource constraint (4) becomes $L_{RD} \leq f(\bar{L}_{RD})$ and production labor resource constraint (5) becomes $L_P \leq 1 - \bar{L}_{RD}$. If $f(L) = aL, a > 0$, the supply of R&D labor is infinitely elastic and R&D wage is constrained to be a multiple of the production wage. If $f$ is strictly concave, one has a model with elastic, but not infinitely elastic, R&D labor supply. In both cases, corporate profits, and hence the incentives for innovation, themselves depend on the amount of innovation in the economy. Computing equilibrium therefore requires finding a fixed point of this feedback loop and, as a result, there is no closed form solution.
3.2 Final goods producer

The final good is produced competitively using production labor and a continuum of intermediate goods \( j \in [0, 1] \) which, at any given time \( t \), exist in a finite set of \( I_{jt} \geq 1 \) qualities \( \{q_{jti}\}_{0 \leq i \leq I_{jt}} \). The production \( Y(t) \) of the final good is given by

\[
Y(t) = F(L_{Ft}, \{q_{jti}\}, \{k_{jti}\}) = \frac{L_{Ft}^\beta}{1-\beta} \int_0^1 \left( \sum_{i=0}^{I_{jt}} q_{jti}^{\frac{\beta}{1-\beta}} k_{jti} \right)^{1-\beta} dj,
\]

where \( k_{jti} \geq 0 \) is the quantity used of intermediate good \( j \) of quality \( q_{jti} \). This specification implies that different qualities of good \( j \) are perfect substitutes. The use of a CES aggregator for intermediate goods with an elasticity that is different from one (i.e., Cobb-Douglas) simplifies the analysis by allowing me to abstract from limit pricing using a simple microfoundation taken from Akcigit and Kerr, 2018. The exponent \( \frac{\beta}{1-\beta} \) on \( q_{jti} \) and later specification of the intermediate goods production and innovation technology together yield balanced growth. The model could be specified with \( q_{jti} k_{jti} \) as the summand instead, a setting which can be interpreted as cost saving innovations. In that case, however, balanced growth requires a similar exponent in the production and innovation technology for intermediate goods. The present specification follows Akcigit and Kerr, 2018. Several alternatives are discussed in Acemoglu, 2009.

Define \( \bar{q}_{jt} = \max_{0 \leq i \leq I_{jt}} \{q_{jti}\} \) as the frontier quality of good \( j \). In equilibrium, the final goods production function admits a simpler representation

\[
Y(t) = F(L_{Ft}, \{\bar{q}_{jt}\}, \{\bar{k}_{jt}\}) = \frac{L_{Ft}^\beta}{1-\beta} \int_0^1 \bar{q}_{jt}^{\frac{\beta}{1-\beta}} \bar{k}_{jt}^{1-\beta} dj.
\]

There is no storage technology for the final good and its price is normalized to 1 in every period.

3.3 Intermediate goods production and innovation

3.3.1 Incumbents

Each quality \( q_{jti} \) of each good \( j \) is produced by a firm which has a monopoly on production of that quality of good \( j \). Intermediate goods \( j \) of any quality are produced according to the production function

\[
k_{jti} = H(\ell_{jti}; Q) = Q \ell_{jti},
\]

where \( \ell_{jti} \geq 0 \) is the labor input and \( Q_t = \int_0^1 \bar{q}_{jt} dj \) is the average frontier quality level in the economy. The producer which produces the frontier quality of good \( j \) is denoted incumbent \( j \). There is no storage of intermediate goods.

As alluded above, the scaling with average quality \( Q_t \) is necessary for a balanced growth path (BGP) in this setting. It implies that improvements in the quality of good \( j \) increases the productivity of all other goods \( j' \neq j \) through a knowledge spillover externality. In this case, balanced growth
requires the scaling to be linear in order to offset the fact that the production wage increases linearly with average quality $Q_t$.

**Intermediate goods market structure** The following setup, which enables me to abstract from limit pricing, is drawn from Akcigit and Kerr, 2018. Within each good $j$, intermediate goods producers play a two-stage Bertrand competition game at each time $t \geq 0$. In the first stage, participants bear a cost of $\varepsilon > 0$ units of the final good in exchange for a right to compete in the second stage. Then, in the second stage, they engage in Bertrand competition against each other and against producers of goods $j' \neq j$. Optimal pricing under Bertrand competition in the second stage implies that all producers not on the frontier will have zero profits. By backward induction, such producers do not pay the $\varepsilon$ entry cost. The incumbent therefore has a second-stage monopoly over good $j$ which implies the optimal price is the monopolistic competition markup $(1 - \beta)^{-1}$ over marginal costs. I study the limit of this model as $\varepsilon \to 0$.

**Own-product innovation** Incumbent $j$ can hire R&D labor to improve the quality of the good she can produce. I refer to this as own-product innovation or OI, following Garcia-Macia et al., 2019 and Klenow and Li, 2020. By performing a flow of $z_{jt}$ units of R&D, she receives a Poisson intensity of $\chi z_{jt}$ of innovating on good $j$, where $\chi > 0$ is an exogenous parameter representing the incumbent’s R&D productivity. This implies an arrival rate of incumbent innovations of

$$\tau_{jt} = \chi z_{jt}. \tag{9}$$

A successful own-product innovation improves the quality of the incumbent’s product by an exogenous step size $\lambda > 1$. Note that incumbent R&D exhibits constant returns to scale. This is necessary to have a closed form solution. When directed at a product of relative quality $\frac{\bar{q}_{jt}}{Q_t}$, the flow cost of $z_{jt}$ units of R&D is $\bar{q}_{jt} z_{jt}$ units of R&D labor. This scaling assumption is natural because higher quality products require more human capital to improve. It implies that there are knowledge spillovers in R&D as an increase in the quality of good $j$ improves the R&D technology of all other goods $j' \neq j$. An assumption of this kind is also necessary for the existence of a balanced growth path, as it ensures that R&D is allocated to all goods $j$ in equilibrium. 

I assume that only incumbents can perform own-product R&D. That is, when an incumbent is overtaken by an entrant or spinout innovation (described in the next section), she loses access to the OI R&D technology and therefore cannot use it to “catch up” to the frontier. One interpretation is that learning by doing means the current producer of a product has unique insights into how to improve on it. This assumption also significantly increases tractability.

---

28 In Appendix E.1 I show that the main result is quantitatively similar when considering a model where incumbents have decreasing returns to scale.
29 The cost of R&D could scale up faster than $\frac{\bar{q}_{jt}}{Q_t}$, which would imply that higher quality products grow slower. In the current setup this would violate BGP since there is no stationary distribution of product quality. Adding a fixed cost, however, would induce such a stationary distribution, because it creates a lower exit barrier. For more discussion of this type of question, see Gabaix, 2009 or Acemoglu and Cao, 2015.
30 Without this assumption, the incumbent problem would have an additional state variable (since falling away from
3.3.2 Spinouts and noncompete agreements

A flow of R&D spending on own-product innovation induces a positive probability of being overtaken by an employee spinout unless the incumbent imposes a noncompete on R&D labor. The decision of whether a noncompete is used is made instant-by-instant and prevents spinouts during the time it is used. To be precise, I assume that if incumbent \( j \) conducts \( z_{jt} \) units of R&D effort, she faces a Poisson intensity of spawning a spinout given by

\[
\tau^S_{jt} = (1 - \mathbb{1}_{j}^{NCA}) \nu z_{jt},
\]

where \( \mathbb{1}_{j}^{NCA} = 1 \) if and only if an NCA is used in that instant. The contract is renegotiated instant-by-instant in the sense that the employee takes as given the incumbent’s NCA policy and wage offered when deciding whether to supply R&D in each instant. Therefore, the incumbent cannot use an NCA without potentially having to offer a higher wage to compensate the employee for the fact that he will not be able to profit from WSO formation.

Upon paying an entry cost (discussed in the next paragraph), an employee spinout from incumbent \( j \) of quality \( \bar{q}_{jt} \) is able to produce good \( j \) with quality \( \lambda \bar{q}_{jt} \). Therefore, it immediately becomes the new incumbent. The previous incumbent’s profits go to zero forever after. The exogenous parameter \( \nu \geq 0 \) encodes the rate at which R&D increases the likelihood of replacement by a WSO. As such, it determines the strength of the key mechanism documented in the preceding empirical section. Recall that \( z_{jt} \) applied to a good of relative quality \( \frac{q}{q_t} \) requires \( \frac{q}{q_t} z_{jt} \) units of R&D labor. Therefore, this specification implies that the rate of spinout generation of a unit of R&D labor is inversely proportional to the relative quality \( \frac{q}{q_t} \) of the good upon which it seeks to innovate. This is consistent with the assumptions made previously about the own-product innovation technology as well as assumptions below regarding the creative destruction innovation technology of entrants.

In order to enter, employee spinouts must pay an entry cost of \( \kappa_e V(j, t | \lambda \bar{q}_{jt}) \) units of the final good in order to begin producing, where \( \kappa_e \in [0, 1) \) is exogenous and \( V(j, t | \lambda \bar{q}_{jt}) \) denotes the equilibrium private value of incumbent \( j \) at time \( t \) with quality \( \lambda \bar{q}_{jt} \). This cost represents non-R&D expenditures required by creative destruction innovation but not by own-product innovation, such as marketing costs or the costs of setting up a new firm.\(^{31}\) The scaling of the cost with \( V(j, t | \lambda \bar{q}_{jt}) \) is natural as it simply means that the cost of entry is proportional to the value of the incumbency

the frontier is no longer an absorbing state) and an additional distribution would need to be tracked (the number of incumbents with the technology to produce each infra-frontier good \( j \)). Typically, papers which focus on catch up growth, such as Aghion et al., 2005, make simplifying assumptions analogous to mine in order to be able to compute the equilibrium in closed form. For a producer \( n \) steps behind the frontier, the assumption of “no catch up innovation” is binding if the expected discounted present value of the cost of \( n + 1 \) innovations using the OI innovation technology is lower than the expected cost of one innovation using the freely available entrant technology (described in Section 3.3.3). For certain parameter values, this inequality will hold for \( n \) sufficiently small.\(^{31}\)

On the other hand, these expenditures could also reflect unmodeled frictions (e.g., asymmetry of information) in the market for ideas which prevent entrants and spinout entrepreneurs from appropriating the full value of their ideas. In this interpretation, entry by spinouts is more socially valuable. Hence, the choice of interpretation does have implications for welfare analysis. The data I will use to calibrate the model will not be able to distinguish between these two interpretations. Therefore, I proceed assuming that they are real costs and consider the transfers interpretation in robustness checks.
position obtained through entry. Further, this scaling assumption allows for an analytical solution to the model.

Incumbents also have to pay a cost in order to use an NCA. Specifically, when incumbent $j$ imposes an NCA on $z_j$ units of R&D, she must pay a flow cost $\kappa_c \nu V(j, t|\bar{q}_j) z_j$ units of the final good. Given (10), incumbent $j$ overall pays

$$NCA \text{ cost}_{jt} = \tau_{jt}^S \kappa_c V(j, t|q). \tag{11}$$

The NCA enforcement cost reflects the direct cost using an NCA. Even if there are no technical restrictions on what kinds of NCAs are valid, determining competition between businesses may be expensive. Moreover, many jurisdictions do, in fact, impose such restrictions, and resources can be invested to prove that the conditions of those restrictions do not apply. Overall, it seems plausible that investing resources increases the likelihood of a successful enforcement of an NCA. Note also that a value of $\kappa_c = \infty$ can be interpreted as a ban on the use of NCAs and a value $\kappa_c = 0$ can be interpreted as a complete relaxation of barriers to the use of NCAs.

As with the entry cost, the factor $V(j, t|q)$ implies that the cost of enforcing NCAs is proportional to the equilibrium value of the incumbent firm. The economic justification is that valuable incumbency positions require more resources to protect via NCAs. In the context of the model, this specification means that the cost of enforcing an NCA on a given unit of R&D labor is proportional to both the value of the WSOs that labor will generate in the absence of an NCA, and the expected loss of incumbent value from an absence of NCAs. As with the entry cost, this assumption also improves model tractability by simplifying the analysis of the optimal noncompete policy (see Section 3.5.3).

Finally, it is important to note that I have assumed that spinout entry does not directly reduce the rate at which incumbent R&D results in successful OI. Instead, spinouts enter when an additional, independent Poisson process with arrival rate $\nu z_j$ has an arrival. The interpretation is that spinouts in this model do not embody stolen ideas that otherwise would have been implemented by the parent firm. Rather, R&D labor generates additional innovations which the employee can use to offer a higher quality product than the incumbent firm. This is consistent with the finding in Klepper, 2007 that spinouts often occur when the employee finds the idea more valuable than his employer rather than through idea stealing. It is important to note that this assumption has important consequences for the private and social usefulness of NCAs. In particular, to the extent that spinouts take ideas that otherwise would be implemented inside the firm, they are less valuable both privately and socially. In the context of this model and calibration, the socially beneficial effects of NCA enforcement would be larger.\footnote{The validity of this assumption could be tested empirically with sufficiently exogenous variation in the enforceability of NCAs or in the use of NCAs. One could study, for example, whether R&D generates fewer patents when R&D managers and employees are not bound by NCAs. However, as noted in the introduction, interpreting the results of such regressions requires care to control for the effects of mobility across enforcement regions or whether the use of NCAs is correlated with firm characteristics in a way that could lead to a spurious correlation.}

20
3.3.3 Entrants

In addition to own-product innovation and spinout formation, growth also results from creative destruction by entrant firms. Specifically, for each frontier quality good $j$ there is a unit mass of entrants indexed by $e \in [0, 1]$. By performing a flow of $\hat{z}_{jet}$ units of R&D, an entrant receives a Poisson intensity of $\hat{z}_{jet} \hat{z}_{jt}^{-\psi}$ of innovating on good $j$, where $\hat{z}_{jt} = \int_0^1 \hat{z}_{jet}de$ total entrant R&D dedicated improving good $j$. From now I drop the bar notation when it is clear from context. Because each entrant is infinitesimal, their R&D efforts have constant returns to scale at the individual level. Note that this means entrant optimization will only pins down $\hat{z}_{jt}$ in equilibrium; as this does not affect aggregates, I will assume that $\hat{z}_{jet} = \hat{z}_{jt}$. Aggregating over $e \in [0, 1]$, the arrival rate of entrant innovations on good $j$ is

$$\hat{\tau}_{jt} = \hat{\chi} \hat{z}_{jt}^{1-\psi}.$$  \hfill (12)

As with the other forms of innovation, the cost in terms of R&D labor is proportional to the relative quality of the good $\bar{q}_{jt} / Q_t$ and a successful innovation yields a monopoly on the production of good $j$ with quality $\lambda \bar{q}_{jt}$.

The parameter $\psi > 0$ introduces decreasing returns at the level of good $j$. It represents a congestion externality in the entrant innovation technology. This is similar to the congestion externality present in search and matching models. Intuitively, due to a lack coordination, entrants attempt similar approaches to solve the same problem. This duplication of effort reduces the overall returns to entrant R&D when considered at the level of good $j$.

Finally, as with spinouts, entrants must also pay an entry cost of $\kappa_e V(j, t|\lambda \bar{q}_{jt})$ in units of the final good in order to enter once they have successfully innovated on a good of quality $\bar{q}_{jt}$. The interpretation is the same as with spinouts as both types of firms are engaging in creative destruction and not own-product innovation.

3.4 Competitive financial intermediary

The representative household owns a competitive financial intermediary which in turn owns all firms in the economy and remits their profits back to the representative household. Individual firms in the economy maximize profits subject to the household’s risk-free discount rate (i.e., there is no collusion due to common ownership). When the representative household receives a shock in good $j$ that allows it to form a spinout, it sells the spinout to the financial intermediary at full private value (i.e. discounting the spinouts profits at the same risk-free discount rate). The financial intermediary takes the entry of the spinout as given, and therefore is willing to pay this value even though the entry of the spinout reduces the value of an existing incumbent.\footnote{As is standard in this type of model, I assume for simplicity that entrants innovate only on frontier quality goods.}

\footnote{The purpose of this construction is to avoid having to assume that the representative household does not take into account the loss of value of the incumbents it owns when spinouts enter.}
3.5 Equilibrium

3.5.1 Definition of equilibrium

The model involves idiosyncratic risk for each good \( j \). In each realization of a given equilibrium, the price at time \( t \) of the frontier intermediate good \( j \), the R&D wage paid by incumbent \( j \), and the allocation of production and R&D labor to the frontier intermediate good \( j \) all depend on the stochastic realization of the frontier quality stochastic process \( \tilde{q}_{jt} \). To study this kind of equilibrium, I look for an equilibrium where prices and quantities are deterministic functions of \( (j, t|\tilde{q}_{jt}) \). I also use \( q \) to refer to \( \tilde{q}_{jt} \) (that is, equilibrium objects are written as functions of \( (j, t|q) \)) when the meaning is clear from context. This approach follows Acemoglu, 2009.

**Definition 1.** An *equilibrium* of this model consists of household consumption \( C(t) \) and bond holdings \( A(t) \); final good production \( Y(t) \); frontier intermediate goods prices \( p(j, t|q) \) and quantities \( k(j, t|q) \); production wages \( \bar{w}(t) \) and production labor allocation to final goods \( L_F(t) \) and intermediate goods \( \ell_I(j, t|q) \); R&D wages paid by entrants \( \hat{w}_{RD}(t) \), by incumbents using and not using noncompetes \( w_{RD}(j, t|q, 1^{NCA}) \); R&D labor allocations across incumbents \( \ell_{RD}(j, t|q) \) and across entrants \( \hat{\ell}_{RD}(j, t|q) \); and noncompete contract allocations \( 1^{NCA}(j, t|q) \) such that

1. The final goods firm maximizes profits.
2. Each incumbent \( j \) optimally chooses production, R&D labor demand, and the use of NCAs.
3. Entrants optimize their R&D labor demand.
4. The representative household optimizes production and R&D labor supply, consumption and savings.
5. The competitive financial intermediary maximizes the discounted present value of profits remitted to the household.

For the sake of tractability, I will restrict attention to symmetric balanced growth path equilibria. This requires two more definitions.

**Definition 2.** A *balanced growth path equilibrium* (BGP) is an equilibrium where there exist \( g, C_0, Q_0 > 0 \) such that

\[
C(t) = C_0 e^{gt}, \quad Q_t = Q_0 e^{gt}.
\]

**Definition 3.** A *symmetric balanced growth path equilibrium* (symmetric BGP) is a BGP where \( z_{jt} = z \) and \( \hat{z}_{jet} = \hat{z} \) for all \( j, e \in [0, 1], t \geq 0 \).
Symmetric BGPs are a natural type of equilibrium given the symmetric setup of the model. It is the typical case studied, e.g. in Grossman and Helpman, 1991 and Acemoglu and Cao, 2015.\(^{35}\)

### 3.5.2 Static equilibrium

The first step is to characterize the static equilibrium given a profile of frontier qualities \(\{\overline{q}_j\}\). Note that I have omitted the dependence on \(t\) for the sake of clarity. In addition, since only the frontier quality is produced in equilibrium, in this section I will drop the \(\overline{q}_j\) notation and refer to the frontier good’s quality and quantity by \(q_j\) and \(k_j\), respectively. The following definition abuses notation slightly as it reuses some of the symbols in the definition of equilibrium.

**Definition 4.** Given a profile a frontier qualities \(\{q_j\}\), a *static equilibrium* consists of frontier intermediate goods prices \(p_j\) and quantities \(k_j\), production wages \(\bar{w}\), and a production labor allocation to final goods \(L_F\) and intermediate goods \(\ell_j\) such that

1. The final goods firm chooses its demand for production labor and intermediate goods in order to maximize profits.

2. Intermediate goods firms choose their demand for production labor in order to maximize profits.

Given this definition, one can directly compute the unique static equilibrium of the model given any profile of qualities. su

**Proposition 1.** Given a profile of frontier qualities \(\{q_j\}\), there exists a unique static equilibrium.

**Proof.** Final goods producer optimization implies the inverse demand function for intermediate goods

\[
p_j = L_F^\beta q_j^\beta k_j^{-\beta}. \tag{13}\]

The intermediate goods market structure implies that incumbent \(j\) can effectively ignore lower quality producers of good \(j\) when making pricing decisions. She therefore sets \(k_j\) to solve

\[
\pi(q_j) = \max_{k_j \geq 0} \left\{ L_F^\beta q_j^\beta k_j^{1-\beta} - \frac{\bar{w}}{Q} k_j \right\}, \tag{14}\]

where \(\bar{w}\) is the equilibrium production goods wage and \(Q = \int_0^1 q_j dj\) is average frontier good quality. The solution to this maximization problem yields expressions for intermediate goods pricing,

\[^{35}\text{One could relax the assumption that } z_{jt} = z \text{ as long as } \int_0^1 z_{jt} \frac{\overline{q}_j}{Q} dj \text{ is constant on the BGP. For a discussion of this, Acemoglu, 2009.}\]
production labor demand, and production, given by

\[ k_j = \left( \frac{(1-\beta)Q}{w} \right)^{1/\beta} L_F q_j, \quad (15) \]

\[ \ell_j = k_j/Q, \quad (16) \]

\[ p_j = \frac{w}{(1-\beta)Q}. \quad (17) \]

Substituting (15) into the first-order condition for final goods firm optimal production labor demand yields an expression for the equilibrium wage \( w \), given by

\[ w = \tilde{\beta} Q, \quad (18) \]

\[ \tilde{\beta} = \beta^\beta (1-\beta)^{1-2\beta}. \quad (19) \]

Substituting (15) and (18) into the expression for profit in (14) yields

\[ \pi_j = (1-\beta)\tilde{\beta} L_F q_j. \quad (20) \]

Substituting (15) into (16) and integrating \( L_I = \int_0^1 \ell_j dj \) yields aggregate labor allocated to intermediate goods production

\[ L_I = \left( \frac{1-\beta}{\beta} \right)^{1/\beta} L_F. \quad (21) \]

Using (21) and \( L_I + L_F = L_P \) in the labor resource constraint (5) yields

\[ L_F = \frac{1 - \bar{L}_{RD}}{1 + \left( \frac{1-\beta}{\beta} \right)^{1/\beta}}. \quad (22) \]

Output can be computed by substituting (15) into (7), yielding

\[ Y = \frac{(1-\beta)^{1-2\beta}}{\beta^{1-\beta}} Q L_F. \quad (23) \]

3.5.3 Existence and uniqueness of symmetric BGP

The static equilibrium derived in the previous section implies, for each profile of qualities \( \{\bar{q}_{jt}\} \), a level of final goods production \( Y(t) \), allocations of production labor, intermediate goods prices, and an allocation of intermediate goods. Moreover, it also implies that the flow of profits to incumbent \( j \) of quality \( q \) is given by \( \pi(j, t|q) = \tilde{\pi} q \), which in turn dictates the dividends the household receives from the financial intermediary. All that remains to characterize the set of symmetric BGPs is to

\[ \Box \]
find R&D wages and labor allocations, consumption, and financial asset holdings that are consistent with household, incumbent, and entrant optimization.

**Household optimization** I begin by deriving the equilibrium conditions that arise from the representative household’s optimization problem. As noted in the definition of equilibrium, the household takes as given incumbent R&D wages and NCA policies $w_{RD}(j,t|q,1_{NCA}), 1_{NCA}(j,t|q)$, entrant R&D wages $\hat{w}_{RD}(t)$, production wages $\bar{w}(t)$, interest rates $\{r_t\}$, and profits from the financial intermediary $\{\Pi_t\}$. To simplify notation, I use the subscript $jt$ to denote a particular realization of an equilibrium object. Specifically, I use $1_{j,t}^{NCA}$ to denote $1^{NCA}(j,t|\bar{q}_{jt}), \ell_{RD,jt}$ to denote $\ell_{RD}(j,t|\bar{q}_{jt},1_{NCA}(j,t|\bar{q}_{jt}))$, $\hat{\ell}_{RD,jt}$ to denote $\hat{\ell}_{RD}(j,t|\bar{q}_{jt})$, and $w_{RD,jt}$ to denote $w_{RD}(j,t|\bar{q}_{jt},1_{j,t}^{NCA})$. I also use $t$ subscripts to denote production wages and the interest rate. Given this, the representative household solves

\[
\begin{align*}
\text{maximize} & \quad \int_0^\infty e^{-\rho t} \frac{C(t)^{1-\theta} - 1}{1 - \theta} \, dt \\
\text{subject to} & \quad \dot{A}(t) = -C(t) + r_t A(t) + \Pi_t + \bar{w}_t L_P(t) \quad \text{(Financial wealth law of motion)} \\
& \quad \quad + \int_0^1 \bar{w}_{RD}(t) \dot{\ell}_{RD,jt} \, dj + \int_0^1 w_{RD,jt} \ell_{RD,jt} \, dj \\
& \quad \quad + \int_0^1 (1 - 1_{j,t}^{NCA}) \nu(\bar{q}_{jt}) (1 - \kappa_e) V(j,t|\lambda \bar{q}_{jt}) \, dj, \\
& \quad A(0) = 0 \quad \text{(Initial wealth)} \\
& \quad \lim_{t \to \infty} e^{-\int_0^t r_s \, ds} A(t) \geq 0, \quad \text{(No Ponzi-game)} \\
& \quad \int_0^1 (\ell_{RD,jt} + \hat{\ell}_{RD,jt}) \, dj \leq \bar{L}_{RD}, \quad \text{(R&D labor endowment)} \\
& \quad L_P(t) \leq 1 - \bar{L}_{RD}. \quad \text{(Production labor endowment)}
\end{align*}
\]

There are also non-negativity constraints on consumption and labor supply. Thus, the household consumes out of profits remitted by the intermediary $\Pi_t$, returns from holdings of the risk free bond $r_t A(t)$, wages earned from production labor supply $\bar{w}_t L(t)$, wages earned from R&D labor supply $\int_0^1 \left(w_{RD,jt} \ell_{RD,jt}(t) + \hat{w}_{RD,\ell_{RD,jt}(t)}\right) \, dj$, and earnings from sales of spinouts to the financial intermediary $\int_0^1 \left(1 - 1_{j,t}^{NCA}\right)(\bar{q}_{jt}) (1 - \kappa_e) V(j,t|\lambda \bar{q}_{jt})) \, dj$.

Because the household’s expected payoff from working at a given incumbent $j$ depends not only on the wage but also on the expected present value of spinouts formed producing good $j$, it is necessary to explicitly model the household’s R&D labor allocation for each individual goods $j$ as a function of its frontier quality realization $\bar{q}_{jt}$. While this is not strictly necessary for R&D labor
supplied to entrants, I present the problem symmetrically.

The expression for the value of earnings from sales of spinouts follows because a spinout occurs in good \( j \) with Poisson intensity \((1 - \frac{1}{NCA(j)}}{\frac{\nu_{RD,j}(t)}{Q(t)}}\) and is sold to the financial intermediary at the price of \((1 - \kappa_e)V(j, t|\lambda q_{jt})\).

36To be more precise, the household forms, and sells, a spinout in each good \( j \) upon the arrival of an independent Poisson process \( N_{jt} \) with time-varying hazard rate \((1 - \frac{1}{NCA(j)}}{\frac{\nu_{RD,j}(t)}{Q(t)}}\). The expected payoff of spinouts from good \( j \) is therefore \((1 - \frac{1}{NCA(j)}}{\frac{\nu_{RD,j}(t)}{Q(t)}})\) \((1 - \kappa_e)V(j, t|\lambda q_{jt})djdt\). The budget constraint integrates over \( j \in [0, 1] \) and, heuristically, divides by \( dt \).

**Lemma 1.** In a symmetric BGP with \( z > 0 \), R&D wages satisfy

\[
\hat{w}_{RD}(t) \leq w_{RD}(j, t|\bar{q}_{jt}) + (1 - \frac{1}{NCA(j, t|\bar{q}_{jt})})\left(\frac{\bar{q}_{jt}}{Q(t)}\right)^{-1}\nu(1 - \kappa_e)V(j, t|\lambda q_{jt})
\]

for all \( t \geq 0 \) and \( j \in [0, 1] \). If \( \hat{z} > 0 \), (24) holds with equality.

**Proof.** Optimality dictates that the household supplies R&D labor only to jobs which provide the highest compensation. Therefore, in order to be consistent with household optimal labor supply when \( z > 0 \), (24) must hold for all \( t \geq 0 \) and \( j \in [0, 1] \). If, in addition \( \hat{z} > 0 \), the household must be indifferent between supplying R&D to incumbents and entrants and (24) must therefore hold with equality. □

Intuitively, in a symmetric BGP, the expected compensation received from each incumbent in exchange for a given amount of R&D labor must be constant. Next, optimal savings gives rise to a standard Euler equation which holds in equilibrium for all \( t \geq 0 \), given by

\[
\frac{\dot{C}(t)}{C(t)} = \frac{1}{\theta}(r_t - \rho).
\]

**Lemma 2.** In a symmetric BGP, \( r_t = r \).

**Proof.** In a symmetric BGP, \( \frac{\dot{C}(t)}{C(t)} = g \) is constant. By (25), \( r_t = \theta g + \rho =: r \). □

Finally, there is a transversality condition on financial wealth. In this model it holds trivially, as \( A(t) = 0 \) in equilibrium. Therefore I omit it from the exposition. The typical transversality condition, which would be imposed here if the household traded claims on the competitive financial intermediary, manifests later in a necessary and sufficient condition for finiteness of household utility.

**Entrant optimization** I briefly consider the optimization problem of entrants in order to prove a result which is useful later. Entrant \( e \) in good \( j \) chooses \( \hat{z}_{jet} \) to maximize profits, taking as given prices (i.e. the R&D wage, the interest rate), good \( j \) entrant R&D \( \hat{z}_{jt} \), and the value of being an incumbent in good \( j \) at time \( t \) with quality \( \lambda q_{jt} \), denoted \( V(j, t|\lambda q_{jt}) \). The resulting optimization
problem is

\[
\max_{\hat{z}_{jt} \geq 0} \hat{\chi}_{jt} \hat{z}_{jt}^{-\psi} (1 - \kappa_e) V(j, t|\lambda \hat{q}_{jt}) - \hat{q}_{jt} \hat{w}_{RD}(t). \tag{26}
\]

\textbf{Lemma 3.} In any equilibrium, \( \hat{z}_{jt} > 0 \). In a symmetric BGP, \( \hat{z} > 0 \).

\textit{Proof.} The first order condition corresponding to (26) is given by

\[
\hat{\chi}_{jt} \hat{z}_{jt}^{-\psi} (1 - \kappa_e) V(j, t|\lambda \hat{q}_{jt}) = \hat{q}_{jt} \hat{w}_{RD}(t).
\]

This requires \( \hat{z}_{jt} > 0 \) because \( Q_t > 0 \) and all other terms are finite in any equilibrium. As \( \forall j \in [0, 1], \hat{z}_{jt} = \hat{z} \) in a symmetric BGP, this implies \( \hat{z} > 0 \). \qed

\textbf{Incumbent optimization} Next, I turn to equilibrium conditions stemming from incumbent and entrant optimization. Incumbent \( j \) takes as given flow profits \( \pi(j, t|q) = \tilde{\pi} q \) derived from the static equilibrium, the interest rate \( r_t \), which she uses to discount future profits, the rate of creative destruction by entrants \( \hat{\tau}(j, t|q) \), as well as the R&D wage she must pay conditional on her choice of NCA policy \( w_{RD}(j, t|q, 1^{NCA}) \). The discounting is at the risk-free rate \( r_t \) because the financial intermediary diversifies across incumbents and there is no aggregate uncertainty.

On the equilibrium path, Lemma 1 gives an expression for the value of the incumbent R&D wage. However, to define the optimal NCA policy of the incumbent, it is necessary to specify the incumbent R&D wage conditional on any choice of NCA policy. Because the R&D labor market is competitive, the incumbent takes as given that she can hire as much R&D labor as necessary provided that she offers total compensation equal to that offered by entrants. If she does not use an NCA, this compensation includes the expected value of spinouts formed by the household.

\textbf{Lemma 4.} In a symmetric BGP, the incumbent maximizes profits taking as given

\[
w_{RD}(j, t|q, 1^{NCA}) + (1 - 1^{NCA})(\frac{q}{Q_t})^{-1} \nu(1 - \kappa_e) V(j, t|q) = \hat{w}_{RD}(t),
\]

\textit{Proof.} By the same argument as in Lemma 1, the expression on the LHS is equal to the expected flow compensation to a unit of R&D labor required to attract any workers. The RHS is the equilibrium price of R&D labor. Lemma 1 already yields the result for the equilibrium choice of \( 1^{NCA} \). The off-equilibrium wage is simply that wage that would be required to actually hire R&D labor in the equilibrium that obtains using the off-equilibrium choice of NCA policy. This is also equal to the equilibrium price of R&D labor \( \hat{w}_{RD}(t) \). \qed

The above discussion implies that the value of an incumbent \( V(j, t|q) \) must satisfy a Hamilton-
The Jacobi-Bellman equation,

\[
(r_t + \bar{\tau}) V(j,t|q) - \dot{V}(j,t|q) = \hat{\pi} q + \max_{1 \leq z \geq 0} \left\{ z \left[ \lambda (V(j,t|\lambda q) - V(j,t|q)) \right] \right\}
\]

where \( \hat{\pi} \) is defined in (20).

The first proposition shows that, in a symmetric BGP, the value function must have a linear form.

**Proposition 2.** *In a symmetric BGP, the value function of the incumbent is given by*

\[
V(j,t|q) = \tilde{V} q,
\]

*for some \( \tilde{V} > 0 \).*

**Proof.** Appendix C.1.1. \( \square \)

The above Proposition implies the following two corollaries.

**Corollary 2.1.** *In a symmetric BGP, there exist \( \tilde{w}_{RD}, \tilde{w}_{RD} > 0 \) such that*

\[
\dot{w}_{RD}(t) = \tilde{w}_{RD} Q_t,
\]

\[
w_{RD}(j,t|q, \mathbb{1}^{NCA}) = \tilde{w}_{RD}(\mathbb{1}^{NCA}) Q_t.
\]

**Proof.** Dividing the entrant first-order condition by \( \bar{q}_{jt} \) yields

\[
\dot{\lambda} \tilde{z}^{\psi} \lambda = \frac{\dot{w}_{RD}(t)}{Q_t}, \tag{28}
\]

implying that \( \dot{w}_{RD}(t)/Q_t \) must be constant, i.e. \( \dot{w}_{RD}(t) = \tilde{w}_{RD} Q_t \) for some \( \tilde{w}_{RD} \). Using this and \( V(j,t|q) = \tilde{V} q \) in Lemma 4 yields \( w_{RD}(j,t|q, \mathbb{1}^{NCA}) = \tilde{w}_{RD}(\mathbb{1}^{NCA}) Q_t \). \( \square \)

From now on, I drop the \( \sim \) superscript when referring to \( \tilde{w}_{RD} \) and \( \tilde{w}_{RD} \) and the meaning is clear from context. The next proposition characterizes the equilibrium NCA policy of all incumbents in a symmetric BGP with \( z > 0 \).
Proposition 3. In a symmetric BGP with $z > 0$, the equilibrium NCA policy of all incumbents is given by

$$
1^NCA_j t = 1^NCA = \begin{cases} 
1 & \text{if } \kappa_c < \bar{\kappa}_c, \\
0 & \text{if } \kappa_c > \bar{\kappa}_c, \\
\{0, 1\} & \text{if } \kappa_c = \bar{\kappa}_c,
\end{cases}
$$

(29)

where $\bar{\kappa}_c = 1 - (1 - \kappa_e)\lambda$.

Proof. Appendix C.1.2.

Proposition 3 implies that, in a symmetric BGP, all incumbents choose the same NCA policy, except on the knife-edge $\kappa_c = \bar{\kappa}_c$ where incumbents are indifferent between $1^NCA = 0$ and $1^NCA = 1$. This results from the fact that the only heterogeneity between incumbents in a symmetric BGP is their quality, which does not affect the optimal policy. The model can be extended to accommodate exogenous heterogeneity in $\kappa_c$ and $\kappa_e$, which would imply a fraction of incumbents would use NCAs. However, there would be no closed form solution and I do not have a clear way to discipline the properties of this distribution.

The proof of Proposition 3 follows from the following equation, which is obtained by substituting into the incumbent HJB the expressions for the equilibrium incumbent R&D wage in Lemma 4.

$$
\begin{align*}
\tilde{r} \tilde{V} &= \tilde{\pi} - \tilde{\tau} \tilde{V} + \max_{1^{NCA} \in \{0, 1\}} \left\{ z \left( \chi(\lambda - 1)\tilde{V} - \tilde{w}_{RD} - (1 - 1^{NCA})(1 - (1 - \kappa_e)\lambda)\nu \tilde{V} - 1^{NCA}\kappa_e \nu \tilde{V} \right) \right\}.
\end{align*}
$$

Equation (30) has an intuitive economic interpretation. The left-hand side is the equilibrium flow payoff on an asset with value $\tilde{V}$. The RHS is the flow payoff of incumbency. The first term, $\tilde{\pi}$, is the flow profits of an incumbent. The term $-\tilde{\tau} \tilde{V}$, is the expected capital loss from creative destruction by entrants. The term $\chi(\lambda - 1)\tilde{V}$ is the expected benefit per unit of R&D effort. The factor $\lambda - 1$ reflects the fact that the incumbent takes into account the opportunity cost of no longer producing with the obsolete technology. The term $-\tilde{w}_{RD}$ reflects the cost of R&D effort due to the contribution from the prevailing R&D wage.

The last two terms determine the optimal NCA policy. The first term $-(1 - 1^{NCA})(1 - (1 - \kappa_e)\lambda)\nu \tilde{V}$ is the expected net cost to the incumbent from not using an NCA. It consists of two components. First, the term $-(1 - 1^{NCA})\nu \tilde{V}$ is the expected capital loss from creative destruction by employee spinouts. Second, the term $(1 - 1^{NCA})(1 - \kappa_e)\lambda \nu \tilde{V}$ is the equilibrium R&D wage discount when an NCA is not imposed. This term is positive because, in equilibrium, the R&D wage
worker accepts a lower wage in return for an expected future payoff from spinout formation, by Lemma 4. Finally, the last term $-\mathbb{1}_{NCA} \kappa c \tilde{V}$ reflects is the direct cost of using an NCA. Incumbents choose the value of $\mathbb{1}_{NCA}$ which maximizes the term multiplying $z$.

**Lemma 5.** In a symmetric BGP with $z > 0$,

$$
\tilde{V} = \frac{\hat{w}_{RD}}{\chi (\lambda - 1) - (1 - \mathbb{1}^{NCA})(1 - (1 - \kappa_e) \lambda) \nu - \mathbb{1}^{NCA} \kappa c \nu}.
$$

(31)

Proof. The constant returns to scale in the incumbent innovation technology implies the incumbent must be indifferent if $z > 0$, i.e. the term multiplying $z$ in (30) must equal zero. Solving for $\tilde{V}$ yields (31). □

**Equilibrium innovation and growth** Finally, I use the results from the previous three sections to characterize the equilibrium R&D allocation and resulting growth rate. Using Proposition 2 and its corollary, the FOC for the entrant’s optimization problem, given in (26), becomes

$$
\hat{\chi} \tilde{z}^{-\psi}(1 - \kappa_e) \lambda \tilde{V} = \hat{w}_{RD}
$$

(32)

Substituting for $\tilde{V}$ by using Lemma 5 yields an expression for entrant R&D,

$$
\tilde{z} = \left(\frac{\hat{\chi}(1 - \kappa_e) \lambda}{\chi (\lambda - 1) - (1 - \mathbb{1}^{NCA})(1 - (1 - \kappa_e) \lambda) \nu - \mathbb{1}^{NCA} \kappa c \nu}\right)^{1/\psi}.
$$

(33)

Market clearing for R&D labor implies

$$
z = \bar{L}_{RD} - \tilde{z}.
$$

(34)

If (34) implies that $z < 0$ then $z = 0$ in equilibrium. I return to this case below. Suppose for now that (34) is consistent with $z > 0$. Then growth is determined by the growth accounting equation\(^{38}\)

$$
g = (\lambda - 1)(\tau + \tau^S + \tilde{\tau}).
$$

(35)

The Euler equation determines the interest rate,

$$
g = \frac{\hat{C}}{C} = \frac{1}{\theta}(r - \rho),
$$

(36)

$$
r = \theta g + \rho.
$$

(37)

Substituting the incumbent’s FOC into the incumbent’s HJB, and using the expression for the interest rate, yields the incumbent’s value $\tilde{V}$,

$$
\tilde{V} = \frac{\tilde{\pi}}{r + \tilde{\tau}}.
$$

(38)

\footnote{A derivation can be found in Appendix C.2.1.}

30
Finally, the entrant optimality condition (32) determines the equilibrium value of \( \hat{w}_{RD} \). If the above steps imply \( z < 0 \) or \( \hat{z} \leq 0 \), then \( z = 0 \) and \( \hat{z} = \hat{L}_{RD} \). The rest of the derivation is the same as in the case that \( z > 0 \).

Finally, in order for equilibrium to be well-defined, household utility must be finite. This is ensured by the following assumption.

**Assumption 1.** \( \rho > (1 - \theta)g \)

In Assumption 1, \( g \) stands for the closed form expression for \( g \). Namely, if \( z > 0 \) then

\[
g = (\lambda - 1)\left( \frac{\hat{\chi}(1 - \kappa_e \lambda)}{\chi(\lambda - 1) - \nu \min\{1 - (1 - \kappa_e)\lambda, \kappa_c\}} \right)^{(1-\psi)/\psi} + \left( \chi + (1 - 1_{\kappa_c < \kappa_e})\nu \right) \left( \hat{L}_{RD} - \frac{\hat{\chi}(1 - \kappa_e \lambda)}{\chi(\lambda - 1) - \nu \min\{1 - (1 - \kappa_e)\lambda, \kappa_c\}} \right)^{1/\psi}. \tag{39}
\]

Otherwise,

\[
g = (\lambda - 1)\hat{\chi}L_{RD}^{1-\psi}. \tag{40}
\]

**Lemma 6.** Under Assumption 1, the household’s utility is finite on a symmetric BGP with growth rate \( g \).

**Proof.** Using \( C(t) = C(0)e^{gt} \) on the BGP, the household’s utility is

\[
U = \mathcal{K} \int_0^\infty e^{-\rho t} e^{(1-\theta)gt} dt + \text{Constant.} \tag{42}
\]

for some constant \( \mathcal{K} > 0 \). The integral \( \int_0^\infty e^{-\rho t} e^{(1-\theta)gt} dt \) converges if and only if \( \rho > (1 - \theta)g \). \( \square \)

Note that \( \theta \geq 1 \) implies Assumption 1. This is the empirically relevant case which I consider in the calibration. The next two assumptions are used in the statement of the main proposition.

**Assumption 2.** \( 0 < \left( \frac{\hat{\chi}(1 - \kappa_e \lambda)}{\chi(\lambda - 1) - (1 - 1_{\kappa_c < \kappa_e})\nu \min\{1 - (1 - \kappa_e)\lambda, \kappa_c\}} \right)^{1/\psi} < \hat{L}_{RD} \).

Assumption 2 ensures that \( z > 0 \) on the symmetric BGP. The term \( 1^NCA \) is as defined in (29). Now I can state the main proposition.

**Proposition 4.** If Assumption 1 holds and \( \kappa_c \neq \kappa_e \) (as defined in Proposition 3), there exists a symmetric BGP. There is a unique symmetric BGP, satisfying \( z > 0 \), if and only if Assumption 2 holds. Otherwise, there exist infinitely many symmetric BGPs that differ only in \( 1^NCA \).

**Proof.** Here I offer a sketch of the proof; details can be found in Appendix C.1.3. Existence follows because a symmetric BGP can be constructed following the derivation given in the main text. When the expression for \( \hat{z} \) is in \((0, \hat{L}_{RD})\), the constructed symmetric BGP has \( z > 0 \) and is unique by construction, as each equilibrium condition has a unique solution given parameters and other equilibrium objects. If \( z = 0 \) then, by the same reasoning, all equilibrium variables are uniquely
determined except for $1_{j^NCA}$. The latter is not pinned down by equilibrium because $z = 0$ implies that the NCA policy is irrelevant to the incumbent.

On the knife-edge $\kappa_c = \bar{\kappa}_c$, the model has a continuum of symmetric BGPs which have different equilibrium outcomes. I discuss these in Appendix C.2.2.

4 Calibration

4.1 Parameters

The model has 11 parameters, $\{\rho, \theta, \beta, \psi, \lambda, \chi, \bar{\chi}, \kappa_e, \kappa_c, \nu, \bar{L}_{RD}\}$. The two parameters $\theta, \psi$ are set externally. The elasticity parameter $\theta$ is set to 2, corresponding to a standard value used in the literature. The entrant R&D curvature parameter $\psi$ is set to a value of 0.5 (the same as in Acemoglu and Cao, 2015), corresponding to a quadratic cost. I consider robustness of the main result to the value of these and all other parameters in Section E.1.

The remaining nine parameters are calibrated to match moments. The parameter $\bar{L}_{RD}$ is calibrated to data on the share of employment in R&D, from the NSF. The parameter $\beta$ is set to match the profit to GDP ratio. The final seven parameters $\{\rho, \lambda, \chi, \bar{\chi}, \kappa_e, \kappa_c, \nu\}$ pertain to preferences ($\rho$) and to the innovation technology and NCA usage (all others), and are chosen to match six moments from the data. One parameter, $\kappa_c$, is partially identified as $\kappa_c > \bar{\kappa}_c$ by the observation that $\tau^S > 0$. The remaining six parameters are exactly identified and the model reproduces the target moments exactly. I discuss the sources of identification in Section 4.3.

4.2 Targets

The targets of the calibration are displayed in Table 3. They consist of the labor productivity growth rate due to creative destruction and own-product innovation, the R&D to GDP ratio, the real return on the corporate sector, the share of growth coming from older firms improving their own products, the employment share of new firms engaging in creative destruction, and the employment share of R&D-induced WSOs. The last three targets are not directly observable in the data. I obtain values for them from previous work in the literature and from the empirical section of this paper.

Matching the productivity growth rate, R&D to GDP ratio and growth share of OI helps calibrate the efficiency of R&D in generating aggregate productivity growth through OI and CD. The real return, profit to GDP ratio and employment share of entering firms determines the discount factor and the reward to innovation. Further, the employment share of young firms helps identify the size of each innovation: at a constant contribution to aggregate growth, a lower share of employment at entering firms implies each innovation by an entering firm is smaller. Finally, matching the employment share of entering WSOs allows the model to capture the rate at which R&D by incumbents increases their likelihood of being replaced by a WSO.

Below, I discuss the calibration targets in more detail.
Growth rate due to CD and OI  The growth rate is calibrated to the growth in labor productivity due to CD and OI, as calculated in Garcia-Macia et al., 2019 and Klenow and Li, 2020. Their accounting procedure uses a structural model of firm dynamics and growth through own-product innovation, creative destruction and new variety creation (but with exogenous productivity growth) to infer the sources of growth from the economy-wide distribution of changes in sales and employment at the firm and establishment levels.\textsuperscript{39}

Growth share of older firms  The growth share of older firms improving their own products is calibrated to the growth share of OI as a fraction of OI and CD innovations, as estimated in Garcia-Macia et al., 2019 and Klenow and Li, 2020. On average they find that, from 1982 to 2013, roughly 65% of CD + OI productivity growth was due to firms at least 6 years old. The computation of the corresponding model moment is described in D.1.3.

R&D spending to GDP  The data on R&D spending is from the National Patterns of R&D resources.\textsuperscript{40} In the data, about half of R&D spending is wages for employees; in the model, the only input to R&D is labor. I opt to match the model’s aggregate R&D intensity to that in the data, including costs other than labor. This means that the model captures the full cost of innovation, but likely overestimates the aggregate labor earnings of R&D employees. The computation of the corresponding model moment is described in D.1.2.

Interest rate  The interest rate in the model corresponds to the discount factor used to price a risky firm. The annual real return on the S&P 500 in the time period 1984-2006 was 9.15%. BAA corporate bonds earned an annual real return of 7.58%. To calculate the excess return of all corporate liabilities, I take a weighted average of the the excess return of stock and equity using the sample period average debt to equity ratio of 58%, or a debt to market value of 37%. This implies an overall real return of 8.57%.

Profits to GDP  The data the profit to GDP ratio comes from the BEA. I use the average ratio during the sample period of 1984-2006.

Employment share of young firms  As discussed in Klenow and Li, 2020, adjustment costs mean that, in the data, it can take several years for a new product to displace an old one. However, in the model, entrants that replace incumbents reach their mature size immediately upon entry, simplifying the model significantly. However, this does mean that, if the model matches the amount of employment in firms of age < 1, it might underestimate the true impact on employment reallocation of each new cohort of firms. For this reason, I choose to target the employment share of firms age \leq 6.

\textsuperscript{39}The key identification assumptions are essentially that creative destruction, own-product innovation and new variety creation all have different implications for the cross-sectional and time-series joint distribution of firm and establishment employment, sales, and exit.

\textsuperscript{40}I take the average of business-funded R&D business-performed R&D.
Table 3: Calibration targets

<table>
<thead>
<tr>
<th>Key parameter(s)</th>
<th>Target</th>
<th>Model</th>
</tr>
</thead>
<tbody>
<tr>
<td>Profit (% GDP)</td>
<td>$\beta$</td>
<td>8.5%</td>
</tr>
<tr>
<td>R&amp;D emp. share</td>
<td>$L_{RD}$</td>
<td>1%</td>
</tr>
<tr>
<td>Real return</td>
<td>$\rho$</td>
<td>8.57%</td>
</tr>
<tr>
<td>Growth rate</td>
<td>$\lambda, \chi, \hat{\chi}$</td>
<td>1.487%</td>
</tr>
<tr>
<td>Age ≥ 6 growth share</td>
<td>$\chi, \hat{\chi}$</td>
<td>65%</td>
</tr>
<tr>
<td>Age &lt; 6 emp. share</td>
<td>$\lambda, \hat{\chi}$</td>
<td>13.34%</td>
</tr>
<tr>
<td>Spinout emp. share</td>
<td>$\nu$</td>
<td>10%</td>
</tr>
<tr>
<td>R&amp;D spending (% GDP)</td>
<td>$\chi, \hat{\chi}, \kappa_e$</td>
<td>1.35%</td>
</tr>
</tbody>
</table>

In addition, I restrict attention only to incumbents and young firms engaging in creative destruction. Because all entry in the model is creative destruction, including employment in entrants developing new varieties would overstate the rate of creative destruction. To do this, I turn to Garcia-Macia et al., 2019 and Klenow and Li, 2020, which estimate the portion of growth coming from firms of different ages engaging in creative destruction, new variety creation, and own product improvement. They find that roughly 18% of employment is in firms age ≤ 6 during the sample period, and that between 30% and 70% of the growth from these firms is due to creative destruction, the rest due to new variety creation. However, in their framework, as in mine, a given amount of growth from creative destruction requires significantly more employment, as it destroys a previous incumbent. Using a value of $\lambda = 1.2$, for example, creative destruction requires 6 times more employment than new variety creation to generate the same amount of growth. Lower values of $\lambda$ imply a larger adjustment. Using $\lambda = 1.2$ as a benchmark value, I calculate an employment share of young firms of 13.34% during the sample period. The calibrated value of $\lambda$ will be 1.084, so a choice of $\lambda = 1.2$ in the calculation of this target is conservative. A lower choice here implies a higher share of employment in young firms and, as discussed in the robustness analysis of Appendix E.1, this would strengthen my main result. The computation of the corresponding model moment is described in D.1.4.

R&D-induced spinout share of employment Finally, matching the employment share of spinouts is of course crucial so that the analysis accurately captures the burden such firms impose on the incumbents that spawn them. As discussed in the empirical section, R&D by parent firms can account for 8.5% of employment at startups in the Venture Source. However, as with the employment share of young firms, I want to restrict attention to firms engaging in creative destruction. The same kind of logic implies an adjustment factor between 1/.93 and 1/.7, which implies that between 9% and 12% of creative destruction startup employment consists of WSOs. I choose a value of 10% for the calibration. The computation of the corresponding model moment is described in D.1.5.
4.3 Identification

In this section, I discuss how the calibration identifies the parameters of the model. First, it is important to note that the relationship between the model parameters and the model-generated moments is non-linear and many parameters influence multiple moments. Figure 3 shows the elasticity of model moments to model parameters. All moments are influenced by multiple parameters. Note that the model is extremely sensitive to the log of $\lambda$. This is in part due to the fact that a proportional change in $\lambda$ implies a much larger proportional change in $\lambda - 1$, which is the percentage by which an innovation increases quality.

The object that is most directly related to identification is the elasticity of inferred model parameters to targeted moments. This is shown in Figure B3. With this perspective, one can see that conclusions based on Figure 3 can be misleading. For example, while an increase in $\lambda$ causes a large increase in the growth share of Age $\geq 6$ firms, increasing the moment target decreases the estimated $\lambda$. Given all the moments that need to be matched, the calibration prefers to match the growth share of Age $\geq 6$ firms by using a much higher $\chi$ and actually slightly lower $\lambda$. To complete the picture, Figure B5 augments Figure 3 with non-calibrated parameters included as both parameters and target moments. As before, Figure B4 inverts this matrix to obtain the elasticity of calibrated parameters to moment targets and non-calibrated parameters. This gives the complete picture of how the model parameters are inferred.

Based on this analysis, one can obtain an intuition regarding how the model parameters are pinned down by the calibration targets. The discount rate $\rho$ is identified to simultaneously match the interest rate, growth rate and IES. The parameters $\beta, \bar{L}_{RD}$ are chosen to exactly match the profit share of GDP and the share of labor in R&D, respectively. The parameters $\lambda, \chi, \hat{\chi}$ are chosen to simultaneously match the growth rate $g$, the share of growth from firms at least six years old, and the employment share of firms less than six years old. Intuitively, these three measures determine the size and frequency of each innovation, as well as its attribution to incumbents or entrants. The parameter $\kappa_e$ is distinguished from $\hat{\chi}$ by matching the R&D to GDP ratio. Intuitively, in order to be consistent with equilibrium, the expected private return to entrant R&D must be the same as the real interest rate. Because entrant R&D has a high expected payoff, as it results in a monopoly, equilibrium requires a large cost. If R&D is the only cost, then the R&D spending to GDP ratio is counterfactually high. Finally, the parameter $\nu$ is identified by matching the employment share of WSOs.

---

41 This is computed as the jacobian matrix of the mapping that takes log parameters to log model moments.

42 This is calculated by inverting the matrix shown in the previous figure. This inverse is locally well-defined because the model is locally exactly identified by the target moments.

43 In this calibration, imposing $\kappa_e = 0$ means that best-fit R&D to GDP ratio is 1.9% rather than 1.35%. This is closely related to the main finding of Comin, 2004.
Table 4: Calibrated parameters

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Value</th>
<th>Description</th>
<th>Source</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\theta$</td>
<td>2</td>
<td>$\theta^{-1} = \text{IES}$</td>
<td>External</td>
</tr>
<tr>
<td>$\psi$</td>
<td>0.5</td>
<td>Entrant R&amp;D curvature</td>
<td>External</td>
</tr>
<tr>
<td>$\rho$</td>
<td>0.056</td>
<td>Discount rate</td>
<td>Internal</td>
</tr>
<tr>
<td>$\beta$</td>
<td>0.094</td>
<td>$\beta^{-1} = \text{Es intermediate goods}$</td>
<td>Internal</td>
</tr>
<tr>
<td>$\lambda$</td>
<td>1.085</td>
<td>Quality ladder step size</td>
<td>Internal</td>
</tr>
<tr>
<td>$\chi$</td>
<td>26.35</td>
<td>Incumbent R&amp;D productivity</td>
<td>Internal</td>
</tr>
<tr>
<td>$\hat{\chi}$</td>
<td>0.461</td>
<td>Entrant R&amp;D productivity</td>
<td>Internal</td>
</tr>
<tr>
<td>$\kappa_e$</td>
<td>0.740</td>
<td>Non-R&amp;D entry cost</td>
<td>Internal</td>
</tr>
<tr>
<td>$\nu$</td>
<td>0.431</td>
<td>Spinout generation rate</td>
<td>Internal</td>
</tr>
<tr>
<td>$\bar{L}_{RD}$</td>
<td>0.01</td>
<td>R&amp;D labor allocation</td>
<td>Internal</td>
</tr>
</tbody>
</table>

Figure 3: Plot showing the elasticity of model moments to model parameters. These elasticities are computed by taking the jacobian matrix of the mapping from log parameters to log model moments.
5 Welfare effect of NCA enforcement and other policies

In this section, I use the calibrated model as a laboratory for studying the effect of policies. I conduct a sequence of second-best analyses assuming the planner can control one or more parameters and/or Pigouvian taxes. All comparisons below are static comparisons between BGPs. Throughout, I assume that taxes (subsidies) are rebated to (financed by) the representative household in a lump-sum payment. Because there is no labor-leisure choice, this does not create any additional distortions in the economy.

The main exercise considers whether, and by how much, reducing barriers to the use of NCAs increases or decreases growth and welfare. Next, I consider a planner who can use R&D subsidies in order to illustrate how they can have an effect in this model even though the total supply of R&D labor is fixed. I then consider a planner who can target R&D subsidies to own-product innovation, a policy which can substitute for the enforcement of NCAs in this setting. Finally, I consider a planner who can simultaneously subsidize own-product innovation and control the enforcement of NCAs.

5.1 Preliminaries

Welfare Social welfare is simply the representative household’s lifetime utility,

\[ \tilde{W} = \int_0^{\infty} e^{-\rho t} C(t)^{1-\theta} \frac{1}{1-\theta} ds. \]  

(43)

Using \( C(t) = \tilde{C} e^{\rho t} \) on the BGP and integrating yields

\[ \tilde{W} = \frac{\tilde{C}^{1-\theta}}{(1-\theta)(\rho - \tilde{g}(1-\theta))} + \kappa(\rho, \theta), \]  

(44)

where \( \kappa(\rho, \theta) \) is a constant that depends only on preferences. Social welfare can thus be decomposed into a growth channel \( (g) \) and an level channel \( (\tilde{C}) \). Higher values for either imply higher welfare. The term \( \tilde{C} \) is referred to as the level of consumption because it is equal to \( C(t)/Q(t) \) on the BGP. It can be further decomposed using

\[ \tilde{C} = \tilde{Y} - (\tilde{\tau} + \tilde{\tau}^S) \tilde{\kappa}_\nu \tilde{V} - \tilde{V}^{NCA} z \tilde{\kappa}_\nu \tilde{V}, \]  

(45)

---

44I do not study the first-best allocation because of the fact that I have specified the model in terms of the private value of being an incumbent, which is only well-defined in a decentralized equilibrium. The model could be augmented – at the expense of requiring numerical solutions – to accomodate the study of a first-best problem by simply specifying the NCA cost and the entry cost as linear in \( \tilde{q}_{jt} \), rather than \( V(j,t|\tilde{q}_{jt}) \).

45A subsidy to intermediate goods revenue or production labor expenditures can also improve welfare by correcting the static inefficiency emanating from the monopolistic competition markup. As this is not the focus of this paper, I do not consider it in the policy analysis.

46This should be written in terms of \( W_t \), the welfare at time \( t \). I ignore this detail in the interest of expository simplicity and without loss of generality since the model grows at constant rate so \( W_t = e^{(1-\theta)\rho t} \tilde{W} \).
so that steady-state consumption is flow output of the final good minus the final goods cost of creative destruction and of NCA enforcement.\footnote{To interpret the entry cost as a transfer, replace (45) with $\check{C} = \check{Y} - 1^{NC}\zeta_{NC}\nu\check{V}$.}

**Consumption-equivalent change in welfare** To make welfare comparisons quantitatively meaningful, I compare welfare across BGP$s in consumption-equivalent (CE) terms. This is defined as follows. For a given equilibrium $\varepsilon$, let $\check{C}_\varepsilon, g_\varepsilon, \check{W}_\varepsilon$ denote the level of consumption given $Q_t$, the growth rate of consumption, and the time-0 present value of household utility, respectively. Following the derivation of welfare in (44), one can write

$$\check{W}_\varepsilon = f(\check{C}_\varepsilon, g_\varepsilon) + \kappa(\rho, \theta), \quad (46)$$

where

$$f(x, y) = \frac{x^{1-\theta}}{(1-\theta)(\rho - y(1-\theta))}, \quad (47)$$

Now consider an equilibrium $\varepsilon'$ such that $\check{W}_\varepsilon < \check{W}_{\varepsilon'}$. The CE welfare improvement from equilibrium $\varepsilon$ to equilibrium $\varepsilon'$ is the permanent increase in consumption in $\varepsilon$ that would achieve the same utility as $\varepsilon'$. More precisely, define $\check{C}_{\varepsilon, \varepsilon'}$ such that

$$f(\check{C}_{\varepsilon, \varepsilon'}, g_\varepsilon) = f(\check{C}_{\varepsilon'}, g_{\varepsilon'}). \quad (48)$$

The CE percentage welfare improvement of $\varepsilon'$ over $\varepsilon$ is the percentage increase from $\check{C}_\varepsilon$ to $\check{C}_{\varepsilon, \varepsilon'}$. For $\theta > 1$ (the case of interest in this paper), a $\frac{\xi}{\theta-1}$% CE welfare improvement results from an $\xi$% decrease in the absolute value of $\check{W} - \kappa(\rho, \theta)$.\footnote{For $\theta < 1$, a $\frac{\xi}{\theta-1}$% CE welfare improvement results from a $\xi$% increase in $\check{W} - \kappa(\rho, \theta)$. The case $\theta = 1$ corresponds to log utility, in which case

$$\check{W} = \frac{\rho \log(\check{C}) + g}{\rho^2} \quad (49)$$

In this case, there is no simple correspondence to obtain CE welfare changes, but they are easy to compute directly. Under the null policy, initial consumption is $\check{C}$ and growth is $g$. Under the new policy, initial consumption is $\check{C}^+$ and growth is $g^+$. The CE welfare change is $\frac{\check{C}^*-\check{C}}{\check{C}}$, where $\check{C}^*$ is defined by

$$\frac{\rho \log(\check{C}^*) + g}{\rho^2} = \frac{\rho \log(\check{C}^+)+g^+}{\rho^2} \quad (50)$$

5.2 NCA cost $\kappa_c$

The first policy I study is the effect of reducing $\kappa_c$. I interpret this as loosening restrictions on the use of NCAs. Note that, as discussed in Section 4, the fact that there are spinouts in the data partially identifies $\kappa_c > \bar{\kappa}_c$. Intuitively, according to the model, the spinouts in the data result from the fact that the NCAs that could have been used to prevent them were either too costly to enforce
Table 5: Effect of reduction in $\kappa_c$ on growth, level of consumption, and welfare

<table>
<thead>
<tr>
<th>Measure</th>
<th>Variable</th>
<th>$\kappa_c &gt; \bar{\kappa}_c$</th>
<th>$\kappa_c = 0$</th>
<th>Chg.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Growth</td>
<td>$g$</td>
<td>1.487%</td>
<td>1.696%</td>
<td>0.21 p.p.</td>
</tr>
<tr>
<td>Level</td>
<td>$\bar{C}$</td>
<td>0.784</td>
<td>0.787</td>
<td>0.39%</td>
</tr>
<tr>
<td>Welfare</td>
<td>$\bar{W}$</td>
<td></td>
<td></td>
<td>3.24% (CE)</td>
</tr>
</tbody>
</table>

Table 6: Decomposition of effect of reducing $\kappa_c$ on growth and R&D

<table>
<thead>
<tr>
<th>Measure</th>
<th>Variable</th>
<th>$\kappa_c &gt; \bar{\kappa}_c$</th>
<th>$\kappa_c = 0$</th>
<th>Chg.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Growth</td>
<td>$g$</td>
<td>1.487%</td>
<td>1.696%</td>
<td>0.21 p.p.</td>
</tr>
<tr>
<td></td>
<td>incumbents</td>
<td>$(\lambda - 1)\tau$</td>
<td>1.20%</td>
<td>1.47%</td>
</tr>
<tr>
<td></td>
<td>entrants</td>
<td>$(\lambda - 1)\hat{\tau}$</td>
<td>0.26%</td>
<td>0.23%</td>
</tr>
<tr>
<td></td>
<td>spinouts</td>
<td>$(\lambda - 1)\tau^S$</td>
<td>0.02%</td>
<td>0%</td>
</tr>
</tbody>
</table>

or legally prohibited. Therefore, to study the effect of reducing barriers to the enforcement of NCAs, I compare the calibrated BGP ($\kappa_c > \bar{\kappa}_c$) to the BGP that obtains when $\kappa_c = 0$. I interpret the latter to a case when there are no legal barriers to the use of NCAs. One might imagine that there is some fundamental cost $\kappa_c$ of using an NCA even in a jurisdiction that imposes no barriers to its use. For simplicity, I assume $\kappa_c = 0$.

The results are displayed in Table 5. Compared to the BGP with $\kappa_c > \bar{\kappa}_c$, the BGP with $\kappa_c = 0$ has higher growth and initial consumption, implying a 3.24% increase in welfare in consumption-equivalent terms. Table 6 shows a decomposition of the sources of growth in the $\kappa_c > \bar{\kappa}_c$ and $\kappa_c = 0$ equilibria. When $\kappa_c = 0$, a higher share of R&D labor is allocated to own-product innovation and there is no longer entry by spinouts. Intuitively, having access to free NCAs ($\kappa_c = 0$) makes own-product innovation effectively less costly than when NCAs are prohibitively expensive ($\kappa_c > \bar{\kappa}_c$) and incumbents face the risk of employee spinout formation. Holding prices constant at their $\kappa_c > \bar{\kappa}_c$ levels, a shift to $\kappa_c = 0$ induces incumbents to demand more R&D labor. To clear the market, the equilibrium compensation to R&D labor must increase. The net effect is a shift in the allocation of R&D labor to own-product innovation.

Figure 4 visually decomposes the effect of reducing $\kappa_c$ on the growth rate and the level of consumption. The top row shows the allocation of R&D and the growth rate that results. For high values of $\kappa_c$ the growth rate is low. As soon as $\kappa_c < \bar{\kappa}_c$, NCAs are used, eliminating spinout entry and reducing the BGP growth rate by a discrete amount. Note that there is no discrete change in the allocation of R&D when $\kappa_c$ crosses this threshold because incumbents are indifferent between using and not using an NCA when $\kappa_c = \bar{\kappa}_c$. As $\kappa_c$ is reduced further, the cost of own-product innovation.
Figure 4: Effect of varying $\kappa_c$ on key equilibrium variables. The top-left panel shows R&D labor allocated to incumbents (own-product innovation) and entrants (creative destruction). The top-right panel shows the aggregate productivity growth rate. The bottom-left panel shows the entry costs paid by entrants and spinouts as well as the direct NCA enforcement cost. Finally, the bottom-right panel shows the level of consumption.

innovation declines (in partial equilibrium), inducing a shift of R&D labor as described in the previous paragraph. Next, the bottom row shows the entry and NCA costs paid and the level of consumption $\tilde{C}$ that results. Plots showing other equilibrium objects (the innovation rate, incumbent value, the interest rate, and the R&D wage) are displayed in Figure B7.

5.2.1 Equilibrium R&D misallocation

To understand how a higher share of R&D labor allocated to own-product innovation increases growth by enough to outweigh the contribution of spinouts, consider the equilibrium marginal effect on innovation (and hence growth) of the two types of R&D. For creative destruction, this is given by

$$\frac{d}{d\tilde{z}} \hat{\tau} = (1 - \psi) \hat{\chi} \hat{z}^{-\psi}.$$  (51)
For own-product innovation, this is given by

\[
\frac{d}{dz} \tau = \chi + (1 - 1^{NCA}) \nu.
\] (52)

A reallocation of R&D labor to own-product innovation increases the BGP growth rate if and only if

\[
\frac{d}{dz} \tau > \frac{d}{dz} \hat{\tau}.
\] (53)

Using (51), (52), and the expression for equilibrium \( \hat{z} \) given in (33), the inequality (53) becomes

\[
1 > \frac{\lambda - 1}{\lambda} \times (1 - \psi) \times \frac{1}{1 - \kappa_e} \times \chi + (1 - 1^{NCA}) \nu \times \frac{\chi(\lambda - 1) - (1 - 1^{NCA})(1 - (1 - \kappa_e) \lambda) \nu - 1^{NCA} \kappa_e \nu}{\chi(\lambda - 1)}.
\] (54)

In the calibrated BGP, the RHS of (54) has a value of 0.13. This means that, in equilibrium, a marginal unit of R&D labor allocated to own-product innovation is 7.9 times as productive in generating growth as a marginal unit of R&D labor allocated to creative destruction. Hence there is a significant amount of misallocation of R&D in equilibrium. This fact is driving the significant growth gains from the R&D reallocation in response to a reduction in \( \kappa_c \). Below I discuss the economic intuition behind each term in (54).

First, the term \( \frac{\lambda - 1}{\lambda} < 1 \) reflects the business stealing externality. Creative destruction by entrants imposes a negative externality on the incumbent. This means that entrants are able to earn the required private return on R&D while experiencing a lower marginal innovation rate per unit of R&D. It is decreasing in \( \lambda \). Intuitively, a lower value of \( \lambda \) means that a higher fraction of the private return to creative destruction is stolen business. Quantitatively, this effect is the main driver of the growth increase from setting \( \kappa_c = 0 \). In the calibration, \( \frac{\lambda - 1}{\lambda} \approx 0.08 \). \(^{49}\)

Next, the term \( 1 - \psi < 1 \) reflects the congestion externality. As discussed in the exposition of the model, individual entrants impose a negative externality on the expected returns of other entrants. This congestion externality leads to an overallocation of R&D to entrants. In the calibration, \( 1 - \psi = 0.5 \). Higher values of \( \psi \) imply stronger congestion externality, exacerbating the equilibrium misallocation of R&D.

The term \( \frac{1}{1 - \kappa_e} \geq 1 \) reflects the additional entry cost paid by entrants upon innovating. It reduces the private payoff from each creative destruction innovation. To ensure that the expected return of creative destruction is the equilibrium rate of return, the expected number of innovations from a given amount of R&D in creative destruction must increase. Therefore, this mechanism works against the previous two effects, and in the calibration, with a value of \( \frac{1}{1 - \kappa_e} \approx 3.84 \), reduces the

\(^{49}\)In models with duopolistic competition between incumbents, such as Aghion et al., 2005, this effect would be attenuated because own-product innovation by a market leader has a business-stealing component.
extent of misallocation of R&D. On the other hand, a higher value of \( \kappa_e \) also reduces the efficiency of creative destruction R&D by increasing the entry cost.\(^{50}\)

Next, the term \( \frac{\chi((1-1NCA)\nu}{\lambda(\lambda-1)(1-(1-\kappa_c)\lambda)\nu-1NCA\kappa_c \nu} \) reflects the contribution to the productivity of own-product innovation stemming from entry by WSOs. If \( 1NCA = 0 \) and \( \nu > 0 \), the term is strictly less than 1. R&D by incumbents has a positive growth externality through its effect on the entry of spinouts. Because these spinouts are not appropriated by the incumbent, in equilibrium a marginal unit of incumbent R&D generates more total innovations. In the calibration, this term is 0.98. This is not very strong as spinout innovations constitute only 1.7% of innovations resulting from incumbent R&D. If \( 1NCA = 1 \) or \( \nu = 0 \) this term is equal to 1.

Finally, the term \( \frac{\chi((\lambda-1)-(1-1NCA)(1-(1-\kappa_e)\lambda)\nu-1NCA\kappa_c \nu}{\lambda(\lambda-1)} \) reflects the fact that entrants pay a different effective cost of R&D than the incumbent. When \( \kappa_c > \tilde{\kappa}_c \), so \( 1NCA = 0 \), the inequality \( 1 - (1 - \kappa_e)\lambda > 0 \) means that the expected harm from future spinouts exceeds the equilibrium wage discount from not using an NCA – i.e., spinouts are bilaterally suboptimal. In equilibrium, this means that incumbents pay a higher effective cost per unit of R&D. When \( \kappa_c < \tilde{\kappa}_c \), so that \( 1NCA = 1 \), incumbents pay the entrant R&D wage plus the NCA enforcement cost; as long as \( \kappa_c > 0 \), incumbents have a strictly higher effective cost of R&D. In equilibrium, the marginal private returns to R&D must equate. Hence, marginal R&D allocated to creative destruction must generate innovations at a lower rate. In the calibration, \( 1NCA = 0 \) and this term is equal to 0.86.\(^{51}\)

**Magnitude of overall effect on growth of reducing \( \kappa_c \)** The preceding discussion shows why there is any scope for a reduction in \( \kappa_c \) to increase growth: inequality (54) implies that the derivative of the growth rate with respect to a reallocation of R&D to own-product innovation is positive. However, the overall magnitude of the increase in growth is the integral of this marginal effect of reallocation minus the reduction growth from the stifling of spinout formation by NCAs. This integral, in turn, is increasing in the sensitivity of the R&D allocation to a given change in \( \kappa_c \nu \), as well as how much \( \kappa_c \nu \) can be reduced from the calibration-implied value of \( \tilde{\kappa}_c \).

The sensitivity of the R&D allocation to a given change in \( \kappa_c \nu \) is determined by the price-elasticity of R&D demand by entrants and incumbents: a higher price-elasticity for either implies a larger R&D reallocation. Note that while a higher price-elasticity of R&D of entrants implies more reallocation, it also implies a weaker congestion externality, reducing the growth impact of a given amount of reallocation of R&D. I consider curvature in the incumbent innovation function in the robustness check in the appendix, discussed briefly in the next subsection.

The extent to which \( \kappa_c \nu \) can be reduced is determined by \( \tilde{\kappa}_c = 1 - (1 - \kappa_e)\lambda \) and \( \nu \). A higher inferred value of \( \tilde{\kappa}_c \) means the shift to \( \kappa_c = 0 \) is a larger reduction in costs. Intuitively, if the calibration finds that creative destruction is more costly per spinout, it implies that there are larger.

\(^{50}\)For the value of \( \kappa_e \) in the calibration, further increasing \( \kappa_e \) on net reduces the benefit to reallocating R&D. This occurs because, for high values of \( \kappa_e \), \( \frac{1}{\sqrt{\kappa_e}} \) becomes highly sensitive to further increases in the value of \( \kappa_e \), while the cost of entry always increases linearly in \( \kappa_e \) (ignoring quantitatively small changes in \( \tilde{V} \)).

\(^{51}\)Alternatively, if \( 1 - (1 - \kappa_e)\lambda < 0 \), spinouts are bilaterally efficient and, in equilibrium, incumbents benefit from spinouts ex ante. As a consequence, incumbents have a lower effective cost of R&D than entrants and the effect is reversed.
barriers to the use of NCAs (a larger value of \( \kappa_c \)). A higher value of \( \nu \) amplifies the beneficial R&D allocation effect of reducing \( \kappa_c \) from \( \bar{\kappa}_c \) to zero as incumbent R&D incentives depend on \( \kappa_c \nu \).

Finally, a higher value of \( \nu \) amplifies the growth decrease from setting \( \kappa_c < \bar{\kappa}_c \) due to the reduction in innovation by WSOs. This loss of growth is equal to \( (\lambda - 1)\nu \bar{z} \), where \( \bar{z} \) is the R&D allocation to own-product innovation in the equilibrium with \( \kappa_c = \bar{\kappa}_c \). Hence, both the growth-increasing and growth-decreasing effects of reducing \( \kappa_c \) are amplified by \( \nu \) in a similar way.

5.2.2 Robustness

I study the robustness of the main result to variation in both target moments and uncalibrated parameters in Appendix E.1. First, I show that the welfare result is quantitatively similar when considering entry costs as transfers to the financial intermediary, at 2.86\% in CE terms. As noted previously, this interpretation increases the social value of within-industry spinouts. The fact that the result is quantitatively similar demonstrates that the key result is not driven by a large resource cost of spinout entry. Second, I show that the result is robust to an augmented model where the incumbent has a decreasing returns to scale innovation production function. The growth increase in that case is similar to the baseline, at 0.177 percentage points, and the welfare increase is also quantitatively similar at 2.65\% in CE terms. Next, I show that the result that setting \( \kappa_c = 0 \) increases growth is robust to a 16\% standard deviation of uncertainty in the calibration targets. Finally, I show that the growth and welfare results are in fact reversed when the model is forced to match a 5\%, rather than 13.34\%, employment share of young firms. This occurs largely due to a higher calibrated value of \( \lambda \) (about 1.59), which works to bring the RHS of inequality (54) much closer to 1. As a result, the improvement in the R&D allocation is not enough to offset the reduction in innovation by spinouts. Growth declines by 0.02 percentage points and welfare decreases by 0.23\% in CE terms.

5.3 R&D subsidy

The first alternative policy I consider is a subsidy to R&D spending. This is a natural class of policy to study due to its significant magnitude the United States, where all told the Federal government funds about 15\% of business-performed R&D.\(^{52}\)

First, I briefly discuss the equilibrium that obtains in the presence of an R&D subsidy; the details of the derivation are in C.3.1. Suppose that the planner subsidizes R&D spending at rate

\(^{52}\)In the United States, the marginal R&D subsidy rate is between 15 and 20\%, which is claimed via deduction on corporate income taxes. The deduction can be carried forward twenty years. These R&D subsidies are applied only to R&D spending above a firm-specific base which is defined in reference to past levels of R&D and firm sales. Taking this into account direct R&D subsidies offer about a 5\% effective subsidy. In addition, federal and local governments directly fund about 10\% of private business-performed R&D.
In this case, in a symmetric BGP the incumbent’s HJB becomes

\[
(r + \tilde{\tau}) \tilde{V} = \tilde{\pi} + \max_{z \geq 0} \left\{ \sum \left( \frac{1}{NCA} \tilde{V} - (1 - T_{RD}) \left( \tilde{w}_{RD} - (1 - 1_{NCA})(1 - \kappa_e)\lambda \nu \tilde{V} \right) - (1 - 1_{NCA})\nu \tilde{V} - 1_{NCA} \kappa_c \nu \tilde{V} \right) \right\}. \tag{55}
\]

Given this HJB, equilibrium usage of NCAs is analogous to the model with no R&D subsidies, although with a different threshold value, given by

\[
\tilde{\kappa}_c = 1 - (1 - T_{RD})(1 - \kappa_e)\lambda. \tag{56}
\]

Note that the threshold is decreasing in the subsidy to R&D spending. This means that a sufficiently large R&D subsidy induces incumbents to use NCAs in equilibrium provided \(\kappa_c < 1\). This is the relevant case, as \(\kappa_c > 1\) means that NCAs are prohibitively expensive even when they require no wage premium in equilibrium.\(^53\)

Using \(z > 0\) and the incumbent FOC as before, the equilibrium R&D allocation is

\[
\hat{z} = \left( \frac{\tilde{\chi}(1 - \kappa_e)\lambda}{\chi(\lambda - 1) - \nu(1_{NCA}\kappa_c + (1 - 1_{NCA})(1 - (1 - T_{RD})(1 - \kappa_e)\lambda))} \right)^{1/\psi}, \tag{57}
\]

\[
z = \tilde{L}_{RD} - \hat{z}. \tag{58}
\]

Inspection of the expression for \(\hat{z}\) shows that, when \(1_{NCA} = 0\), \(\hat{z}\) increases in \(T_{RD}\). Intuitively, an increase in the R&D subsidy from \(T_{RD}^0\) to \(T_{RD}^1\) reduces the wage expenses paid for R&D by the factor \(1 - \frac{1 - T_{RD}^0}{1 - T_{RD}^1}\) for both incumbents and entrants. However, the incumbent’s effective cost of R&D also includes the increased likelihood of creative destruction by an employee spinout, which is not subsidized. Therefore, the incumbent’s effective cost of R&D is reduced by a factor \(\tilde{\tau}_{RD} < 1 - \frac{1 - T_{RD}^1}{1 - T_{RD}^0}\).

In general equilibrium, an increase in \(T_{RD}\) results in R&D labor being reallocated to entrants. Furthermore, for large enough subsidies, \(\kappa_c < \tilde{\kappa}_c\) so the incumbent uses NCAs. This occurs because she prefers to pay employees using subsidized wages rather than through future employee spinouts, the cost of which – i.e., the expected lost profit from being replaced by a spinout – is not subsidized.

Turning to the quantitative exercise, recall that the model can only set identify \(\kappa > \tilde{\kappa}_c\) in order to match the fact that there are spinouts. However, if \(\kappa_c\) is much larger than \(\tilde{\kappa}_c\), any changes to the incentives for NCAs induced by policy will have no observable effect on the equilibrium. In order to be able to illustrate these effects, in this and all subsequent exercises I assume that \(\kappa_c = 1.1\tilde{\kappa}_c\).\(^54\)

\(^{53}\) The case \(\kappa_c = 1\) means the incumbent is indifferent about NCAs when the R&D is fully subsidized.

\(^{54}\) In a model with incumbent heterogeneity, such as non-degenerate distribution of \(\kappa_e\) and \(\kappa_c\) across goods \(j\), one could avoid making such a stark assumption by identifying certain moments of this distribution by matching the fraction of workers who are bound by NCAs. Still, some assumptions on the form of this distribution would likely be necessary, as a single balanced growth path cannot does not characterize it.
Table 7: Effect of R&D subsidy on growth, NCAs, level of consumption and welfare

<table>
<thead>
<tr>
<th>Measure</th>
<th>Variable</th>
<th>R&amp;D Subsidy (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>0</td>
</tr>
<tr>
<td>Growth</td>
<td>g</td>
<td>1.49%</td>
</tr>
<tr>
<td>Level</td>
<td>$\bar{C}$</td>
<td>0.784</td>
</tr>
<tr>
<td>NCAs 1$^{NCA}$</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Δ Welfare (CE) $\bar{W}$</td>
<td>$-0.18%$</td>
<td>$-0.36%$</td>
</tr>
</tbody>
</table>

Table 8: Decomposition of effect of R&D subsidies on growth and R&D

<table>
<thead>
<tr>
<th>Measure</th>
<th>Variable</th>
<th>R&amp;D Subsidy (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>0</td>
</tr>
<tr>
<td>Growth</td>
<td>$g$</td>
<td>1.49%</td>
</tr>
<tr>
<td>incumbents</td>
<td>$(\lambda - 1)\tau$</td>
<td>1.20%</td>
</tr>
<tr>
<td>entrants</td>
<td>$(\lambda - 1)\hat{\tau}$</td>
<td>0.26%</td>
</tr>
<tr>
<td>spinouts</td>
<td>$(\lambda - 1)\tau^S$</td>
<td>0.02%</td>
</tr>
<tr>
<td>R&amp;D</td>
<td>$z/\bar{L}_{RD}$</td>
<td>54.0%</td>
</tr>
<tr>
<td>incumbents (%)</td>
<td>$\hat{z}/\bar{L}_{RD}$</td>
<td>46.0%</td>
</tr>
</tbody>
</table>

Table 7 shows how the equilibrium growth rate $g$, consumption level $\bar{C}$, usage of NCAs $1^{NCA}$, and welfare $\bar{W}$ are affected by R&D subsidies. The growth rate declines and the level of consumption is unaffected. When R&D subsidies are raised to 30%, incumbents use NCAs. Welfare declines overall at an increasing rate, particularly when NCAs are used and spinouts no longer enter. Table 8 decomposes the sources of growth and allocation of R&D. As in the case of varying $\kappa_c$, the growth decrease results from a reallocation of R&D. Because the reallocation is to entrants (creative destruction) and away from incumbents (own-product innovation) in this case, the growth rate decreases due to (54). Figure 5 shows how key equilibrium objects respond to R&D subsidies. In particular, growth declines continuously with the R&D subsidy until point at which $\kappa_c < \bar{\kappa}_c$, where there is a discrete fall in the growth rate as spinouts no longer enter. Plots describing how the entire equilibrium responds are displayed in Figure B8.

Because the supply of R&D labor is inelastic in this model, it is not surprising that untargeted R&D subsidies are do not lead to higher growth rates. However, in a model without within-industry spinouts, untargeted R&D would simply have no effect on the equilibrium. This exercise shows that the general equilibrium adjustment of R&D wages can lead to a counterproductive misallocation of R&D in response to R&D subsidies. To the extent that the supply of R&D labor is not perfectly elastic, this mechanism can play a role. It implies that untargeted R&D subsidies have a weaker positive effect (less “bang for buck”) than in a standard model without within-industry spinouts and...
Figure 5: Summary of equilibrium for baseline parameter values and various values of $T_{RD}$. This assumes that $\kappa_c = 1.1\bar{\kappa}_c$. The top-left panel shows R&D labor allocated to incumbents (own-product innovation) and entrants (creative destruction). The top-right panel shows the aggregate productivity growth rate. The bottom-left panel shows the entry costs paid by entrants and spinouts as well as the direct NCA enforcement cost. Finally, the bottom-right panel shows the level of consumption.
NCAs. This mechanism will apply to the extent that firms with less to lose from spinout formation engage in inefficient forms of R&D.

5.4 Targeted R&D subsidy

The fact that reducing \( \kappa_c \) increases growth by reallocating R&D to own-product innovation suggests that targeted R&D subsidies can be a substitute to enforcing NCAs. Suppose that the planner can subsidize own-product R&D while excluding creative destruction R&D, denoting this subsidy \( T_{RD,I} < 0 \) (tax if \( T_{RD,I} < 0 \)). The new equilibrium conditions are derived in Appendix C.3.2. Here I simply discuss the equilibrium use of NCAs and the R&D allocation. As before, NCAs are used in equilibrium as long as \( \kappa_c \) is lower than a threshold, this time given by

\[
\hat{\kappa}_c = 1 - (1 - T_{RD,I})(1 - \kappa_e)\lambda.
\]  

Therefore, just as with untargeted R&D subsidies, large enough targeted R&D subsidies have the unintended effect of inducing the use of NCAs. The equilibrium R&D allocation is given by

\[
\hat{z} = \left( \frac{(1 - T_{RD,I})\hat{\chi}(1 - \kappa_e)\lambda}{\chi(\lambda - 1) - \nu(1 - \kappa_e)} \right)^{1/\psi},
\]

\[
z = \bar{L}_{RD} - \hat{z}.
\]

Note that when \( 1^{NCA} = 0 \), the term \( T_{RD,I} \) appears in both the numerator and denominator with the same negative sign. Specifically, the numerator has the term \(-\hat{\chi}(1 - \kappa_e)\lambda T_{RD,I}\) and the denominator has the term \(-(1 - 1^{NCA})\nu(1 - \kappa_e)\lambda T_{RD,I} \). The term in the numerator results from the fact that incumbents demand more R&D labor in response to subsidies, which in equilibrium increases the wage and thus reduces incentives for entrants to do R&D. The term in the denominator is analogous to the term in the case of untargeted R&D subsidies. It results from the fact that when \( 1^{NCA} = 0 \), incumbents effectively do not receive the full R&D subsidy because part of their cost of R&D is the expected loss of business due to employee spinout formation. Hence, the second effect dampens the first effect, but does not overturn it.

To see this formally, differentiate (60) with respect to \( T_{RD,I} \) and simplify, yielding

\[
\frac{d\hat{z}^\psi}{dT_{RD,I}} = \frac{-\chi(\lambda - 1) + \nu}{g(T_{RD,I}, \chi, \lambda, \kappa_c, \kappa_e)^2},
\]

where \( g \) is a function of \( T_{RD,I} \) and certain model parameters. As \( g^2 > 0 \), it follows that \( \frac{d\hat{z}^\psi}{dT_{RD,I}} \) is strictly negative if and only if

\[
\chi(\lambda - 1) > \nu.
\]

The condition (63) holds in the calibration. In fact, it is also a necessary condition for \( \tau^S > 0 \) in
To quantify the effect of a targeted R&D subsidy, I return to the calibrated model. Table 9 shows how the equilibrium growth rate $g$, consumption level $\tilde{C}$, and welfare are affected by subsidies to incumbent R&D. The growth rate and level of consumption increase. When R&D subsidies are raised to 40%, incumbents are induced to use NCAs. Welfare increases at a decreasing rate, particularly when NCAs are used and spinouts no longer enter. Table 10 decomposes the effect on growth and R&D. As targeted R&D subsidies are increased, R&D is reallocated to incumbents. Recalling the derivation above, this occurs because $\chi(\lambda - 1) - \nu > 0$. Because inequality (54) holds, this increases growth.

Figure 6 plots the effect of targeted R&D subsidies on the R&D allocation, growth rate, entry and NCA costs, and the level of consumption. Note how there is a discrete decline in the growth rate for targeted R&D subsidies above 25% due to the use of NCAs they induce. Consumption increases in the targeted R&D subsidy due to a decline in entry costs. When NCAs begin to be used in equilibrium, there is a very small discrete fall in consumption as the rising NCA cost is largely offset by a declining cost of entry by spinouts. Additional plots showing how other equilibrium objects are affected by targeted R&D subsidies can be found in Figure B9.

Before turning to the optimal policy analysis, note that in this model a tax on R&D allocated...
to creative destruction would serve a very similar role to the targeted R&D subsidy studied in this section. In fact, denoting the tax by $T_{RD,E}$, the R&D allocation in that case would be

$$\hat{z} = \left( \frac{(1 + T_{RD,E})^{-1} \hat{\chi}(1 - \kappa_e)\lambda}{\hat{\chi}(\lambda - 1) - \nu(\mathbb{1}_{NCA}\kappa_e + (1 - \mathbb{1}_{NCA})(1 - (1 - \kappa_e)\lambda))} \right)^{1/\psi}, \quad (64)$$

$$z = \bar{L}_{RD} - \hat{z}. \quad (65)$$

In this case, the R&D allocation is more sensitive because there is no dampening secondary effect and there is no change in the NCA usage threshold $\hat{\kappa}_c$. I have chosen to focus on the subsidy to incumbent R&D for two reasons. First, it may be impossible to implement a tax on R&D expenditures in reality as a firm can simply claim that it is not conducting R&D. The regulator would need to check whether firms are engaging in creative destruction R&D. Contrast that with the case of R&D subsidies, which require the regulator only to evaluate the validity of firms’ claimed R&D expenses. Second, the equivalence of these two policies would no longer hold in a model with an elastic total supply of R&D labor. In that case, a tax on R&D, even if targeted, would reduce the total amount of R&D.

### 5.5 NCA cost $\kappa_c$ and targeted R&D subsidy

While the targeted R&D subsidy analyzed in the last section can increase growth and welfare significantly, it also induces the use of NCAs, dampening its positive effect on growth and welfare. This suggests that the optimal policy involves targeted R&D subsidies and an increase in the cost of NCA enforcement. Therefore, in this section I consider a planner who can simultaneously control $\kappa_c$ and $T_{RD,I}$.

Table 11 shows the effect on growth of various combinations of $\kappa_c$ and $T_{RD,I}$. For low values of $T_{RD,I}$, growth is maximized when $\kappa_c = 0$, as in the baseline case. As $T_{RD,I}$ increases, the benefit of reducing $\kappa_c$ declines. In this sense, targeted R&D subsidies substitute for the enforcement of NCAs. With targeted R&D subsidies at 40%, the growth increase from setting $\kappa_c = 0$ is less than one third of its value without subsidies. At 60% subsidies, the growth increase from $\kappa_c = 0$ is reduced to 0.01 percentage points. For intermediate values of $\kappa_c$, the growth rate is slightly lower. When targeted R&D subsidies are at 80%, reducing $\kappa_c$ to zero reduces the growth rate. The threshold occurs at a targeted R&D subsidy of 62% (not shown in the table). Growth is maximized at 2.282% when the targeted R&D subsidy is 88% and NCAs are banned.

Table 12 shows the CE welfare percent improvement relative to the baseline calibration. The same results are shown in a contour plot in Figure 7. Compared to growth, enforcing NCAs remains socially optimal for higher values of the targeted R&D subsidy. This results from the reduced entry costs. In the table, enforcing NCAs reduces welfare only with an 80% targeted subsidy; the specific threshold in this case is a targeted subsidy of 77%. The peak improvement in welfare is achieved with a high targeted R&D subsidy of around 95% and a ban on NCAs. The welfare increase resulting from this policy is 11.44% in CE terms. One caveat is that the welfare gain from banning NCAs is
Figure 6: Summary of equilibrium for baseline parameter values and various values of $T_{RD,I}$. This assumes that $\kappa_c = 1.1\bar{\kappa}_c$. The top-left panel shows R&D labor allocated to incumbents (own-product innovation) and entrants (creative destruction). The top-right panel shows the aggregate productivity growth rate. The bottom-left panel shows the entry costs paid by entrants and spinouts as well as the direct NCA enforcement cost. Finally, the bottom-right panel shows the level of consumption.
Table 11: Growth rate for various values of \( \kappa_c \) and \( T_{RD,I} \)

<table>
<thead>
<tr>
<th>( T_{RD,I} ) (%)</th>
<th>( \kappa_c/\bar{\kappa}_c )</th>
<th>0</th>
<th>.5</th>
<th>1</th>
<th>1.5</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>1.70%</td>
<td>1.59%</td>
<td>1.49%</td>
<td>1.49%</td>
<td></td>
</tr>
<tr>
<td>20</td>
<td>1.92%</td>
<td>1.86%</td>
<td>1.78%</td>
<td>1.80%</td>
<td></td>
</tr>
<tr>
<td>40</td>
<td>2.09%</td>
<td>2.05%</td>
<td>2.02%</td>
<td>2.03%</td>
<td></td>
</tr>
<tr>
<td>60</td>
<td>2.20%</td>
<td>2.19%</td>
<td>2.17%</td>
<td>2.19%</td>
<td></td>
</tr>
<tr>
<td>80</td>
<td>2.24%</td>
<td>2.24%</td>
<td>2.24%</td>
<td>2.28%</td>
<td></td>
</tr>
<tr>
<td>90</td>
<td>2.24%</td>
<td>2.24%</td>
<td>2.24%</td>
<td>2.28%</td>
<td></td>
</tr>
</tbody>
</table>

Table 12: Welfare improvement from varying \( \kappa_c \) and targeted R&D subsidy

<table>
<thead>
<tr>
<th>( T_{RD,I} ) (%)</th>
<th>( \kappa_c/\bar{\kappa}_c )</th>
<th>0</th>
<th>.5</th>
<th>1</th>
<th>1.5</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>3.24%</td>
<td>1.70%</td>
<td>0%</td>
<td>0%</td>
<td></td>
</tr>
<tr>
<td>20</td>
<td>6.47%</td>
<td>5.53%</td>
<td>4.35%</td>
<td>4.47%</td>
<td></td>
</tr>
<tr>
<td>40</td>
<td>8.79%</td>
<td>8.25%</td>
<td>7.60%</td>
<td>7.74%</td>
<td></td>
</tr>
<tr>
<td>60</td>
<td>10.31%</td>
<td>10.03%</td>
<td>9.71%</td>
<td>9.95%</td>
<td></td>
</tr>
<tr>
<td>80</td>
<td>11.12%</td>
<td>10.98%</td>
<td>10.82%</td>
<td>11.17%</td>
<td></td>
</tr>
<tr>
<td>90</td>
<td>11.27%</td>
<td>11.16%</td>
<td>11.04%</td>
<td>11.42%</td>
<td></td>
</tr>
</tbody>
</table>

relatively small at only about 0.15 percentage points in CE terms.

In Appendix E.2 I consider how this result changes when interpreting the entry costs of creative destruction as transfers to the financial intermediary. The finding that the growth and welfare optimal policies involve large targeted R&D subsidies and a ban on NCAs still holds. However, the optimal targeted R&D subsidy is lower, at about 88%, and the threshold targeted R&D subsidy beyond which banning NCAs is socially optimal drops from 78% to 62%. Moreover, the gain from banning NCAs in the optimal policy is three times larger, at 0.45 percentage points. Intuitively, treating entry costs as transfers makes creative destruction by entrants and spinouts more socially valuable. Hence, the optimal allocation involves less reallocation of R&D towards own-product innovation and the benefit from banning NCAs and therefore allowing spinout formation is greater.
Figure 7: Summary of equilibrium for baseline parameter values and various values of $T_{RD,I}$ and $\kappa_c$. The optimal policy improves welfare by 11.44% in consumption-equivalent terms.
6 Conclusion

This paper studies the effect of NCAs on growth and welfare. It makes several contributions. Empirically, it shows that R&D tends to predict employee spinout formation at the firm level, after controlling for firm level variables and various fixed effects. The relationship is statistically and economically significant, accounting for about 8.5% of employment in the Venture Source data. It also finds that WSOs are on average 35% larger than equivalent non-WSO startups in terms of employment, revenue, and valuation per founder. These results are consistent with previous findings in Babina and Howell, 2019 and Muendler et al., 2012, respectively, in a dataset of publicly traded parent firms and venture-capital financed startups. This is important as publicly traded firms conduct the majority of R&D and VC-funded startups are particularly important contributors to aggregate productivity growth.

Theoretically, it extends a textbook model of endogenous growth with creative destruction to allow for within-industry spinouts and noncompete agreements. The augmented model offers some new theoretical insights about the effects of noncompete enforcement on growth and welfare. When calibrated to the firm-level relationship between R&D and employee spinouts, findings from the growth accounting literature, and standard aggregate data, it suggests that reducing barriers to the usage of NCAs can significantly increase growth and welfare. It reaches this conclusion even in a model where there is a fixed supply of the input to R&D due to an inferred equilibrium misallocation of R&D labor to creative destruction, in turn resulting from strong inferred business stealing and congestion externalities of creative destruction.

The analysis also finds some novel implications of policies aimed at increasing growth. R&D subsidies may reduce growth by shifting R&D labor to firms engaging in creative destruction. They also induce the use of NCAs. R&D subsidies targeted at own-product innovation improve the allocation of R&D but continue to induced the socially inefficient use of NCAs. Hence, it is optimal to combine targeted subsidies to own-product innovation with a ban on the use of NCAs.

There are several possibilities for further work in this direction. On the empirical side, it is important to study the extent to which within-industry spinouts, as identified here, compete with their parent firms. Further, it could be fruitful to analyze the relationship between R&D spending and WSO formation separately by industry, as there is evidence for this kind of heterogeneity in my dataset. Finally, one could explore how NCA enforceability, or the actual presence of an NCA, affects the level of R&D spending as well as the relationship between R&D and the formation of WSOs.

The model could be extended in several ways at the cost of requiring a numerical solution. As mentioned in the text, the aggregate supply of R&D labor could be made price-elastic so that the total amount of R&D responds to incentives. One could allow for heterogeneity in $\kappa_c$ and $\kappa_e$, which would yield a fraction of incumbents using NCAs in equilibrium. The model could then be disciplined with data on the prevalence of noncompetes found in Starr et al., 2019. Finally, one could set up the model to have limit pricing and varying markups. This would allow confronting
data on how markups and concentration vary by NCA enforceability.\textsuperscript{56}

\textsuperscript{56}This is the approach taken in Baslandze, 2019. I discuss how my work differs from hers in the literature review of the introduction.
# Appendix of tables

<table>
<thead>
<tr>
<th>Industry</th>
<th>Startups</th>
<th>Individuals</th>
<th>State</th>
<th>Startups</th>
<th>Individuals</th>
<th>Year</th>
<th>Startups</th>
<th>Individuals</th>
</tr>
</thead>
<tbody>
<tr>
<td>Business Applications Software</td>
<td>1890</td>
<td>24572</td>
<td>California</td>
<td>8921</td>
<td>110208</td>
<td>1987</td>
<td>353</td>
<td>2680</td>
</tr>
<tr>
<td>Biotechnology Therapeutics</td>
<td>1089</td>
<td>14381</td>
<td>Massachusetts</td>
<td>2279</td>
<td>30536</td>
<td>1988</td>
<td>356</td>
<td>2831</td>
</tr>
<tr>
<td>Communications Software</td>
<td>1036</td>
<td>13387</td>
<td>New York</td>
<td>1644</td>
<td>16896</td>
<td>1989</td>
<td>403</td>
<td>3247</td>
</tr>
<tr>
<td>Advertising/Marketing</td>
<td>962</td>
<td>12004</td>
<td>Texas</td>
<td>1372</td>
<td>15524</td>
<td>1990</td>
<td>396</td>
<td>3159</td>
</tr>
<tr>
<td>Network/Systems Management Software</td>
<td>688</td>
<td>10866</td>
<td>Pennsylvania</td>
<td>927</td>
<td>8906</td>
<td>1991</td>
<td>422</td>
<td>3751</td>
</tr>
<tr>
<td>Vertical Market Applications Software</td>
<td>561</td>
<td>7065</td>
<td>Washington</td>
<td>827</td>
<td>9647</td>
<td>1992</td>
<td>537</td>
<td>4854</td>
</tr>
<tr>
<td>Online Communities</td>
<td>550</td>
<td>3984</td>
<td>Colorado</td>
<td>637</td>
<td>7592</td>
<td>1993</td>
<td>554</td>
<td>5294</td>
</tr>
<tr>
<td>Application-Specific Integrated Circuits</td>
<td>464</td>
<td>6202</td>
<td>Virginia</td>
<td>618</td>
<td>7670</td>
<td>1994</td>
<td>689</td>
<td>6735</td>
</tr>
<tr>
<td>IT Consulting</td>
<td>461</td>
<td>5482</td>
<td>Georgia</td>
<td>579</td>
<td>6420</td>
<td>1995</td>
<td>876</td>
<td>8910</td>
</tr>
<tr>
<td>Wired Communications Equipment</td>
<td>454</td>
<td>6651</td>
<td>Illinois</td>
<td>573</td>
<td>5981</td>
<td>1996</td>
<td>1191</td>
<td>13102</td>
</tr>
<tr>
<td>Drug Development Technologies</td>
<td>412</td>
<td>5038</td>
<td>New Jersey</td>
<td>567</td>
<td>6560</td>
<td>1997</td>
<td>1141</td>
<td>13426</td>
</tr>
<tr>
<td>Healthcare Administration Software</td>
<td>403</td>
<td>4978</td>
<td>Florida</td>
<td>559</td>
<td>5277</td>
<td>1998</td>
<td>1513</td>
<td>19471</td>
</tr>
<tr>
<td>Therapeutic Devices (Minimally Invasive/Noninvasive)</td>
<td>374</td>
<td>4668</td>
<td>North Carolina</td>
<td>466</td>
<td>5326</td>
<td>1999</td>
<td>2557</td>
<td>32463</td>
</tr>
<tr>
<td>Fiberoptic Equipment</td>
<td>364</td>
<td>4911</td>
<td>Maryland</td>
<td>437</td>
<td>5230</td>
<td>2000</td>
<td>2003</td>
<td>24251</td>
</tr>
<tr>
<td>Database Software</td>
<td>357</td>
<td>4500</td>
<td>Minnesota</td>
<td>378</td>
<td>4093</td>
<td>2001</td>
<td>1067</td>
<td>13268</td>
</tr>
<tr>
<td>Business Support Services: Other</td>
<td>341</td>
<td>3889</td>
<td>Ohio</td>
<td>375</td>
<td>2911</td>
<td>2002</td>
<td>986</td>
<td>12928</td>
</tr>
<tr>
<td>Multimedia/Streaming Software</td>
<td>337</td>
<td>4066</td>
<td>Connecticut</td>
<td>367</td>
<td>3727</td>
<td>2003</td>
<td>1037</td>
<td>11912</td>
</tr>
<tr>
<td>Entertainment</td>
<td>335</td>
<td>2555</td>
<td>Utah</td>
<td>255</td>
<td>2570</td>
<td>2004</td>
<td>1110</td>
<td>13345</td>
</tr>
<tr>
<td>Procurement/Supply Chain</td>
<td>327</td>
<td>4744</td>
<td>Oregon</td>
<td>224</td>
<td>2415</td>
<td>2005</td>
<td>1222</td>
<td>13305</td>
</tr>
<tr>
<td>Wireless Communications Equipment</td>
<td>322</td>
<td>4506</td>
<td>Tennessee</td>
<td>222</td>
<td>2134</td>
<td>2006</td>
<td>1380</td>
<td>13823</td>
</tr>
<tr>
<td>Specialty Retailers</td>
<td>308</td>
<td>2915</td>
<td>Arizona</td>
<td>211</td>
<td>2280</td>
<td>2007</td>
<td>1506</td>
<td>13045</td>
</tr>
<tr>
<td>Pharmaceuticals</td>
<td>301</td>
<td>3629</td>
<td>Michigan</td>
<td>209</td>
<td>1707</td>
<td>2008</td>
<td>1416</td>
<td>10469</td>
</tr>
<tr>
<td>Data Management Services</td>
<td>292</td>
<td>4094</td>
<td>Wisconsin</td>
<td>142</td>
<td>1110</td>
<td>2009</td>
<td>1494</td>
<td>9460</td>
</tr>
</tbody>
</table>

Table A1: Statistics on startups covered by VS sample. Industry information uses VS industrial classification. Startups are counted by founding year, individuals by year they joined the firm.
<table>
<thead>
<tr>
<th>Title</th>
<th>Individuals</th>
<th>Percentage</th>
</tr>
</thead>
<tbody>
<tr>
<td>Chief Executive Officer</td>
<td>11154</td>
<td>25.3</td>
</tr>
<tr>
<td>Chief Technology Officer</td>
<td>8493</td>
<td>19.2</td>
</tr>
<tr>
<td>President &amp; CEO</td>
<td>7965</td>
<td>18.0</td>
</tr>
<tr>
<td>Chief</td>
<td>5893</td>
<td>13.3</td>
</tr>
<tr>
<td>President</td>
<td>4118</td>
<td>9.3</td>
</tr>
<tr>
<td>Founder</td>
<td>3000</td>
<td>6.8</td>
</tr>
<tr>
<td>Chairman &amp; CEO</td>
<td>2439</td>
<td>5.5</td>
</tr>
<tr>
<td>President &amp; COO</td>
<td>982</td>
<td>2.2</td>
</tr>
<tr>
<td>President &amp; Chairman</td>
<td>105</td>
<td>0.2</td>
</tr>
</tbody>
</table>

Table A2: Most frequent titles among key founders in VS data.

<table>
<thead>
<tr>
<th>Year</th>
<th>Number of founders</th>
<th>Number of start-ups</th>
<th>Number of founders from public companies</th>
<th>Fraction from public companies (%)</th>
<th>Fraction from public companies when bio. info available (%)</th>
<th>Fraction from public companies in same 4-digit NAICS (%)</th>
<th>Fraction from public companies in same 4-digit NAICS when bio. info available (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1986</td>
<td>269</td>
<td>216</td>
<td>45</td>
<td>16.7</td>
<td>22.8</td>
<td>5.2</td>
<td>7.1</td>
</tr>
<tr>
<td>1987</td>
<td>356</td>
<td>280</td>
<td>43</td>
<td>12.1</td>
<td>15.1</td>
<td>3.9</td>
<td>4.9</td>
</tr>
<tr>
<td>1988</td>
<td>372</td>
<td>281</td>
<td>58</td>
<td>15.6</td>
<td>19.9</td>
<td>4.6</td>
<td>5.8</td>
</tr>
<tr>
<td>1989</td>
<td>479</td>
<td>341</td>
<td>75</td>
<td>15.7</td>
<td>19.2</td>
<td>4.2</td>
<td>5.1</td>
</tr>
<tr>
<td>1990</td>
<td>478</td>
<td>329</td>
<td>85</td>
<td>17.8</td>
<td>21.1</td>
<td>6.3</td>
<td>7.5</td>
</tr>
<tr>
<td>1991</td>
<td>540</td>
<td>356</td>
<td>81</td>
<td>15.0</td>
<td>17.9</td>
<td>6.3</td>
<td>7.5</td>
</tr>
<tr>
<td>1992</td>
<td>674</td>
<td>450</td>
<td>100</td>
<td>14.8</td>
<td>17.9</td>
<td>3.3</td>
<td>3.9</td>
</tr>
<tr>
<td>1993</td>
<td>778</td>
<td>490</td>
<td>137</td>
<td>17.6</td>
<td>20.3</td>
<td>6.7</td>
<td>7.7</td>
</tr>
<tr>
<td>1994</td>
<td>999</td>
<td>611</td>
<td>167</td>
<td>16.7</td>
<td>19.3</td>
<td>4.9</td>
<td>5.7</td>
</tr>
<tr>
<td>1995</td>
<td>1326</td>
<td>772</td>
<td>224</td>
<td>16.9</td>
<td>19.0</td>
<td>5.2</td>
<td>5.8</td>
</tr>
<tr>
<td>1996</td>
<td>1926</td>
<td>1077</td>
<td>319</td>
<td>16.6</td>
<td>18.1</td>
<td>4.9</td>
<td>5.3</td>
</tr>
<tr>
<td>1997</td>
<td>1986</td>
<td>1036</td>
<td>345</td>
<td>17.4</td>
<td>19.0</td>
<td>5.9</td>
<td>6.5</td>
</tr>
<tr>
<td>1998</td>
<td>2895</td>
<td>1390</td>
<td>541</td>
<td>18.7</td>
<td>19.6</td>
<td>5.2</td>
<td>5.5</td>
</tr>
<tr>
<td>1999</td>
<td>5189</td>
<td>2388</td>
<td>975</td>
<td>18.8</td>
<td>19.6</td>
<td>5.0</td>
<td>5.2</td>
</tr>
<tr>
<td>2000</td>
<td>4084</td>
<td>1832</td>
<td>786</td>
<td>19.2</td>
<td>20.4</td>
<td>5.2</td>
<td>5.5</td>
</tr>
<tr>
<td>2001</td>
<td>2235</td>
<td>948</td>
<td>384</td>
<td>17.1</td>
<td>18.7</td>
<td>6.3</td>
<td>6.9</td>
</tr>
<tr>
<td>2002</td>
<td>2113</td>
<td>884</td>
<td>385</td>
<td>18.2</td>
<td>20.1</td>
<td>7.3</td>
<td>8.0</td>
</tr>
<tr>
<td>2003</td>
<td>1979</td>
<td>903</td>
<td>344</td>
<td>17.4</td>
<td>19.8</td>
<td>7.5</td>
<td>8.5</td>
</tr>
<tr>
<td>2004</td>
<td>2098</td>
<td>988</td>
<td>365</td>
<td>17.4</td>
<td>20.1</td>
<td>6.8</td>
<td>7.9</td>
</tr>
<tr>
<td>2005</td>
<td>2278</td>
<td>1068</td>
<td>400</td>
<td>17.6</td>
<td>20.7</td>
<td>6.5</td>
<td>7.7</td>
</tr>
<tr>
<td>2006</td>
<td>2492</td>
<td>1212</td>
<td>432</td>
<td>17.3</td>
<td>20.5</td>
<td>6.3</td>
<td>7.5</td>
</tr>
<tr>
<td>2007</td>
<td>2817</td>
<td>1366</td>
<td>388</td>
<td>13.8</td>
<td>17.0</td>
<td>4.9</td>
<td>6.1</td>
</tr>
<tr>
<td>2008</td>
<td>2710</td>
<td>1307</td>
<td>422</td>
<td>15.6</td>
<td>19.1</td>
<td>5.4</td>
<td>6.6</td>
</tr>
</tbody>
</table>

Table A3: Summary of founders. Here, "founder" includes all individuals employed at startups in the VentureSource database who (1) joined the startup within 3 year(s) of its founding year; and (2) have the title of CEO, CTO, CCEO, PCEO, PRE, PCHM, PCOO, FDR, CHF.
### Table A4: Robustness of PPML regression without controls to adding industry-age, industry-year, and state-year fixed effects. Clustering is at the firm level to ensure sufficient clusters.

<table>
<thead>
<tr>
<th></th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
</tr>
</thead>
<tbody>
<tr>
<td>log(R&amp;d)</td>
<td>0.40***</td>
<td>0.52***</td>
<td>0.44***</td>
<td>0.47***</td>
</tr>
<tr>
<td></td>
<td>(0.044)</td>
<td>(0.067)</td>
<td>(0.062)</td>
<td>(0.072)</td>
</tr>
<tr>
<td>Firm FE</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Industry-Age FE</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Industry-Year FE</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>State-Year FE</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
</tr>
<tr>
<td>Clustering</td>
<td>Firm</td>
<td>Firm</td>
<td>Firm</td>
<td>Firm</td>
</tr>
<tr>
<td>pseudo R-squared</td>
<td>0.28</td>
<td>0.31</td>
<td>0.33</td>
<td>0.34</td>
</tr>
<tr>
<td>Observations</td>
<td>5308</td>
<td>4928</td>
<td>4784</td>
<td>4254</td>
</tr>
</tbody>
</table>

Standard errors in parentheses

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table A4: Robustness of PPML regression without controls to adding industry-age, industry-year, and state-year fixed effects. Clustering is at the firm level to ensure sufficient clusters.

### Table A5: Robustness of PPML regression with controls to adding industry-age, industry-year, and state-year fixed effects. Clustering is at the firm level to ensure sufficient clusters.

<table>
<thead>
<tr>
<th></th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
</tr>
</thead>
<tbody>
<tr>
<td>log(R&amp;d)</td>
<td>0.50</td>
<td>0.51**</td>
<td>0.83***</td>
<td>2.17**</td>
</tr>
<tr>
<td></td>
<td>(0.34)</td>
<td>(0.26)</td>
<td>(0.29)</td>
<td>(0.90)</td>
</tr>
<tr>
<td>Firm FE</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Age FE</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Industry-Year FE</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>State-Year FE</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
</tr>
<tr>
<td>Clustering</td>
<td>Firm</td>
<td>Firm</td>
<td>Firm</td>
<td>Firm</td>
</tr>
<tr>
<td>pseudo R-squared</td>
<td>0.33</td>
<td>0.35</td>
<td>0.35</td>
<td>0.36</td>
</tr>
<tr>
<td>Observations</td>
<td>630</td>
<td>625</td>
<td>471</td>
<td>279</td>
</tr>
</tbody>
</table>

Standard errors in parentheses

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table A5: Robustness of PPML regression with controls to adding industry-age, industry-year, and state-year fixed effects. Clustering is at the firm level to ensure sufficient clusters.
Table A6: Dependent variable is the logarithm of the number of employees while the independent variable is the fraction of founders who most recently worked at a public firm in the same industry. The first column shows the raw regression. The following three columns control for state, industry, time, cohort and age factors. Specifically, each regression uses a subset of two of the three (year, age, cohort) effects. Further, these fixed effects are (separately) interacted with State and Industry. I use multiway clustering at the State and Industry levels.

<table>
<thead>
<tr>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
</tr>
</thead>
<tbody>
<tr>
<td>WSO4 founders</td>
<td>0.20</td>
<td>0.32***</td>
<td>0.32***</td>
</tr>
<tr>
<td>Total founders</td>
<td>(0.18)</td>
<td>(0.023)</td>
<td>(0.025)</td>
</tr>
<tr>
<td>State-Year FE</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>State-Age FE</td>
<td>No</td>
<td>Yes</td>
<td>No</td>
</tr>
<tr>
<td>State-Cohort FE</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
</tr>
<tr>
<td>NAICS4-Year FE</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>NAICS4-Age FE</td>
<td>No</td>
<td>Yes</td>
<td>No</td>
</tr>
<tr>
<td>NAICS4-Cohort FE</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
</tr>
<tr>
<td>Clustering</td>
<td>State, Industry</td>
<td>State, Industry</td>
<td>State, Industry</td>
</tr>
<tr>
<td>R-squared (adj.)</td>
<td>0.00063</td>
<td>0.38</td>
<td>0.41</td>
</tr>
<tr>
<td>R-squared (within, adj)</td>
<td>0.00063</td>
<td>0.0025</td>
<td>0.0025</td>
</tr>
<tr>
<td>Observations</td>
<td>54424</td>
<td>53509</td>
<td>53311</td>
</tr>
</tbody>
</table>

Standard errors in parentheses
* p < 0.1, ** p < 0.05, *** p < 0.01

Table A7: Dependent variable is the logarithm of annual revenue while the independent variable is the fraction of founders who most recently worked at a public firm in the same industry. The first column shows the raw regression. The following three columns control for state, industry, time, cohort and age factors. Specifically, each regression uses a subset of two of the three (year, age, cohort) effects. Further, these fixed effects are (separately) interacted with State and Industry. I use multiway clustering at the State and Industry levels.

<table>
<thead>
<tr>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
</tr>
</thead>
<tbody>
<tr>
<td>WSO4 founders</td>
<td>-0.13</td>
<td>0.45***</td>
<td>0.42***</td>
</tr>
<tr>
<td>Total founders</td>
<td>(0.094)</td>
<td>(0.13)</td>
<td>(0.081)</td>
</tr>
<tr>
<td>State-Year FE</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>State-Age FE</td>
<td>No</td>
<td>Yes</td>
<td>No</td>
</tr>
<tr>
<td>State-Cohort FE</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
</tr>
<tr>
<td>NAICS4-Year FE</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>NAICS4-Age FE</td>
<td>No</td>
<td>Yes</td>
<td>No</td>
</tr>
<tr>
<td>NAICS4-Cohort FE</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
</tr>
<tr>
<td>Clustering</td>
<td>State, Industry</td>
<td>State, Industry</td>
<td>State, Industry</td>
</tr>
<tr>
<td>R-squared (adj.)</td>
<td>0.000092</td>
<td>0.30</td>
<td>0.38</td>
</tr>
<tr>
<td>R-squared (within, adj)</td>
<td>0.000092</td>
<td>0.0030</td>
<td>0.0026</td>
</tr>
<tr>
<td>Observations</td>
<td>16948</td>
<td>15500</td>
<td>15531</td>
</tr>
</tbody>
</table>

Standard errors in parentheses
* p < 0.1, ** p < 0.05, *** p < 0.01
Table A8: Dependent variable is the logarithm of post-money valuation while the independent variable is the fraction of founders who most recently worked at a public firm in the same industry. The first column shows the raw regression. The following three columns control for state, industry, time, cohort and age factors. Specifically, each regression uses a subset of two of the three (year, age, cohort) effects. Further, these fixed effects are (separately) interacted with State and Industry. I use multiway clustering at the State and Industry levels.

<table>
<thead>
<tr>
<th></th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
</tr>
</thead>
<tbody>
<tr>
<td>WSO4 founders</td>
<td>0.46***</td>
<td>0.42***</td>
<td>0.36***</td>
<td>0.33***</td>
</tr>
<tr>
<td>Total founders</td>
<td>(0.065)</td>
<td>(0.058)</td>
<td>(0.069)</td>
<td>(0.074)</td>
</tr>
<tr>
<td>State-Year FE</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
<td>No</td>
</tr>
<tr>
<td>State-Age FE</td>
<td>No</td>
<td>Yes</td>
<td>No</td>
<td>Yes</td>
</tr>
<tr>
<td>State-Cohort FE</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>NAICS4-Year FE</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
<td>No</td>
</tr>
<tr>
<td>NAICS4-Age FE</td>
<td>No</td>
<td>Yes</td>
<td>No</td>
<td>Yes</td>
</tr>
<tr>
<td>NAICS4-Cohort FE</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
</tr>
</tbody>
</table>

Clustering | State, Industry | State, Industry | State, Industry | State, Industry |
R-squared (adj.) | 0.0042 | 0.28 | 0.29 | 0.26 |
R-squared (within, adj) | 0.0042 | 0.0050 | 0.0035 | 0.0028 |
Observations | 26504 | 25174 | 25027 | 25337 |

Standard errors in parentheses
* p < 0.1, ** p < 0.05, *** p < 0.01

Table A9: Dependent variable is the annual hazard rate (%) of going out of business while the independent variable is the fraction of founders who most recently worked at a public firm in the same industry. The first column shows the raw regression. The following three columns control for state, industry, time, cohort and age factors. Specifically, each regression uses a subset of two of the three (year, age, cohort) effects. Further, these fixed effects are (separately) interacted with State and Industry. I use multiway clustering at the State and Industry levels.

<table>
<thead>
<tr>
<th></th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
</tr>
</thead>
<tbody>
<tr>
<td>WSO4 founders</td>
<td>-0.17*</td>
<td>-0.57****</td>
<td>-0.52****</td>
<td>-0.54****</td>
</tr>
<tr>
<td>Total founders</td>
<td>(0.091)</td>
<td>(0.095)</td>
<td>(0.11)</td>
<td>(0.10)</td>
</tr>
<tr>
<td>State-Year FE</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
<td>No</td>
</tr>
<tr>
<td>State-Age FE</td>
<td>No</td>
<td>Yes</td>
<td>No</td>
<td>Yes</td>
</tr>
<tr>
<td>State-Cohort FE</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>NAICS4-Year FE</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
<td>No</td>
</tr>
<tr>
<td>NAICS4-Age FE</td>
<td>No</td>
<td>Yes</td>
<td>No</td>
<td>Yes</td>
</tr>
<tr>
<td>NAICS4-Cohort FE</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
</tr>
</tbody>
</table>

Clustering | State, Industry | State, Industry | State, Industry | State, Industry |
R-squared (adj.) | 0.0000013 | 0.031 | 0.033 | 0.018 |
R-squared (within, adj) | 0.0000013 | 0.000057 | 0.000043 | 0.000046 |
Observations | 240155 | 239696 | 239788 | 239959 |

Standard errors in parentheses
* p < 0.1, ** p < 0.05, *** p < 0.01

59
Table A10: Dependent variable is the annual hazard rate (%) of a successful exit (IPO, acquisition) while the independent variable is the fraction of founders who most recently worked at a public firm in the same industry. The first column shows the raw regression. The following three columns control for state, industry, time, cohort and age factors. Specifically, each regression uses a subset of two of the three (year, age, cohort) effects. Further, these fixed effects are (separately) interacted with State and Industry. I use multiway clustering at the State and Industry levels.

<table>
<thead>
<tr>
<th>Clustering</th>
<th>State, Industry</th>
<th>State, Industry</th>
<th>State, Industry</th>
<th>State, Industry</th>
</tr>
</thead>
<tbody>
<tr>
<td>R-squared (adj.)</td>
<td>0.00046</td>
<td>0.035</td>
<td>0.033</td>
<td>0.035</td>
</tr>
<tr>
<td>R-squared (within, adj)</td>
<td>0.00046</td>
<td>0.00034</td>
<td>0.00027</td>
<td>0.00027</td>
</tr>
<tr>
<td>Observations</td>
<td>240155</td>
<td>239696</td>
<td>239788</td>
<td>239959</td>
</tr>
</tbody>
</table>

Standard errors in parentheses
* p < 0.1, ** p < 0.05, *** p < 0.01

Table A11: 2-digit NAICS codes summary

<table>
<thead>
<tr>
<th>NAICS Code</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>11</td>
<td>Agriculture, Forestry, Fishing and Hunting</td>
</tr>
<tr>
<td>21</td>
<td>Mining, Quarrying, and Oil and Gas Extraction</td>
</tr>
<tr>
<td>22</td>
<td>Utilities</td>
</tr>
<tr>
<td>23</td>
<td>Construction</td>
</tr>
<tr>
<td>31-33</td>
<td>Manufacturing</td>
</tr>
<tr>
<td>42</td>
<td>Wholesale trade</td>
</tr>
<tr>
<td>44-45</td>
<td>Retail trade</td>
</tr>
<tr>
<td>48-49</td>
<td>Transportation and warehousing</td>
</tr>
<tr>
<td>51</td>
<td>Information</td>
</tr>
<tr>
<td>52</td>
<td>Finance and insurance</td>
</tr>
<tr>
<td>53</td>
<td>Real estate and Rental and Leasing</td>
</tr>
<tr>
<td>54</td>
<td>Professional, Scientific, and Technical Services</td>
</tr>
<tr>
<td>55</td>
<td>Management of Companies and Enterprises</td>
</tr>
<tr>
<td>56</td>
<td>Administrative, Support, Waste Management, Remediation Service</td>
</tr>
<tr>
<td>61</td>
<td>Educational services</td>
</tr>
<tr>
<td>62</td>
<td>Health Care and Social Assistance</td>
</tr>
<tr>
<td>71</td>
<td>Arts, Entertainment, Recreation</td>
</tr>
<tr>
<td>72</td>
<td>Accomodation and Food Services</td>
</tr>
<tr>
<td>81</td>
<td>Other Services (except public Admin.)</td>
</tr>
<tr>
<td>92</td>
<td>Public Administration</td>
</tr>
</tbody>
</table>

Table A11: 2-digit NAICS codes summary
Figure B1: Heatmap displaying the distribution of parent 2-digit NAICS code (row), conditional on child NAICS code (column). Darker hues indicate a higher density. The industries 49 and 99 are omitted because I do not link any startups in those industries to parent firms in Compustat.
Figure B2: Economic magnitude of regression estimates (i.e. using the average value of the coefficient on R&D from the first two columns of Table 2). The figure compares predictions for WSO founder counts at the industry-year level (4-digit industries) to the observed number of founders. This plot considers only industries starting with 3 (manufacturing) or 5 (which includes software publishers and professional services), which together comprise more than 85% of R&D spending in Compustat. The points are color coded by 1-digit NAICS industry. The line in the corresponding color shows the fit of a regression of predicted founders on actual founders. The dotted black line shows the 45 degree line.
Figure B3: Plot showing the elasticity of parameters to moments. It is computed by inverting the jacobian matrix of the mapping from log parameters to log model moments (whose entries comprise the previous figure). These elasticities, along with estimates of the noisiness of the moments used in the calibration, can be used to estimate confidence intervals for the parameters in the model, and thereby for the welfare comparison in question. On each subplot, the horizontal axis labels refer, from left to right, to the interest rate, the growth rate, the share of growth in firms age > 6 (I for incumbent), the share of growth in firms age < 6 (E for entrant), the share of employment in spinouts, and the R&D to GDP ratio.
Figure B4: Same as Figure B3, but now including non-calibrated parameters. As before, this calculated by inverting the jacobian displayed in Figure B5. On each subplot, the horizontal axis labels refer, from left to right, to the interest rate, the growth rate, the share of growth in firms age > 6 (I for incumbent), the share of growth in firms age < 6 (E for entrant), the share of employment in spinouts, and the R&D to GDP ratio.
Figure B5: Plot showing the elasticity of moments to model parameters, including parameters chosen externally $\theta, \beta, \psi$. These non-calibrated parameters are added in as effective moments to be matched, allowing the sensitivity of calibrated parameters $\rho, \lambda, \bar{\chi}, \kappa_E, \nu$ to these parameters to be computed by simply inverting this matrix, as before.
Figure B6: Sensitivity of welfare comparison to model parameters. This is $\nabla_p \log (CE \text{ Welfare chg.%) }$, where $W(p)$ maps log parameters to the log of the percentage change in BGP consumption which is equivalent to the change in welfare from changing $\kappa_c$ from $\kappa_c > \bar{\kappa}_c$ to $\kappa_c = 0$ (i.e. going from banning to frictionlessly enforcing NCAs).
Figure B7: Effect of varying $\kappa_c$ on equilibrium variables and welfare.
Figure B8: Summary of equilibrium for baseline parameter values and various values of the untargeted R&D subsidy $T_{RD}$. This assumes that $\kappa = 1.1\kappa_c$. 
Figure B9: Summary of equilibrium for baseline parameter values and various values of the targeted R&D subsidy $T_{RD,I}$. This assumes that $\kappa_c = 1.1\bar{\kappa}_c$. 
C  Model appendix

C.1  Proofs of propositions

C.1.1  Proof of Proposition 2

Below, I assume that the value of incumbency in line $j$ is differentiable in $t$ conditional on there being no innovations. This excludes by assumption equilibria in which there may be stochastic jumps in the value function. To my knowledge, this restriction is standard in treatments of this type of model, such as in the seminal Grossman and Helpman, 1991 and Acemoglu, 2009.

Proof. If $z > 0$ then the FOC holds with equality; otherwise, we can ignore that terms multiplied by $z$ in the incumbent’s HJB. Hence, the incumbent HJB implies

$$(r + \hat{\tau})V(j, t|q) - \dot{V}(j, t|q) = \tilde{\pi}q,$$  \hspace{1cm} (66)

where I used $\hat{\tau}_{jt} = \hat{\tau}$ on a symmetric BGP. This differential equation has a time-independent solution, given by

$$V(j, t|q) = \frac{\tilde{\pi}}{r + \hat{\tau}}q$$  \hspace{1cm} (67)

$$= \tilde{V}q.$$  \hspace{1cm} (68)

The incumbent HJB also has time-dependent solutions. Intuitively, one has to rule out equilibria where the wage to consumption ratio shrinks or grows exponentially proportionally with $V(j, t|q)$.

To prove this, first rearrange the time-dependent incumbent HJB to obtain

$$\dot{V}(j, t|q) = (r + \hat{\tau})V(j, t|q) - \tilde{\pi}q.$$  \hspace{1cm} (69)

Next, differentiate (69) to obtain

$$\ddot{V}(j, t|q) = (r + \hat{\tau})\dot{V}(j, t|q).$$  \hspace{1cm} (70)

This means that if $\dot{V} < 0$ initially, then $\dot{V} < 0$ initially as well, and similarly if $\dot{V} > 0$ initially then $\dot{V} > 0$ initially as well. Hence, if in equilibrium $V(j, t|q)$ drifts locally, it must drift monotonically.

Suppose first that $\dot{V}(j, t|q) < 0$. Then (70) implies that $\dot{V}(j, t|q)$ tends to $-\tilde{\pi}q$ as $t \to \infty$. This means that $V$ reaches a negative value in finite time with positive probability. This contradicts optimality since the incumbent is always free to choose $z = 0$ and earn flow profits $\tilde{\pi}q$.

Next, suppose that $\dot{V} > 0$ initially, then

$$\dot{V} = r + \hat{\tau} - \frac{\tilde{\pi}q}{V},$$  \hspace{1cm} (71)

so $V$ grows at a rate that asymptotically approaches the exponential rate $r + \hat{\tau}$. Suppose first that
Differentiating the FOC of the incumbent FOC yields
\[
-\frac{\dot{V}(j, t|q)}{V(j, t|q)} = g - \frac{\dot{w}_{RD}(t)}{\dot{w}_{RD}(t)}.
\] (72)

This implies that with positive probability \( \dot{w}_{RD}(t) \dot{z} > \dot{Y} \), which is incompatible with final goods market clearing.

If instead \( z = 0 \), taking logs and differentiating the entrant optimality condition yields
\[
\frac{\dot{V}(j, t|\lambda q)}{V(j, t|\lambda q)} = \frac{\dot{w}_{RD}(t)}{\dot{w}_{RD}(t)} - g,
\] (73)

where I used \( \frac{d \log(x)}{dt} = \frac{\dot{x}}{x} \) and \( \dot{Q}_t = g \). This is the same as (72) but evaluated at the quality one step above the frontier quality. It means that if the incumbent value is growing for constant \( q \), the R&D wage must grow faster than the economy. This yields the same contradiction as before.

### C.1.2 Proof of Proposition 3

**Proof.** Using the representation \( V(j, t|q) = \tilde{V} q \) derived in Proposition 2 in the incumbent HJB (27) and dividing both sides by \( q \) yields
\[
(r + \dot{\tau}) \tilde{V} = \tilde{\pi} + \max_{1^{NCA} \in \{0, 1\}} \left\{ z \left( \chi(\lambda - 1) \tilde{V} - w_{RD}(1^{NCA}) - (1 - 1^{NCA})\nu \tilde{V} - 1^{NCA}\kappa_c \nu \tilde{V} \right) \right\}.
\]

In any symmetric BGP with \( z > 0 \), Lemma 1 determines the relationship between \( w_{RD,jt}(1^{NCA}) \) and \( \dot{w}_{RD,t} \). Substituting in \( w_{RD}(1^{NCA}) \) using the indifference condition (24) derived in Lemma 1 yields
\[
(r + \dot{\tau}) \tilde{V} = \tilde{\pi} + \max_{1^{NCA} \in \{0, 1\}} \left\{ z \left( \chi(\lambda - 1) \tilde{V} - w_{RD} - (1 - 1^{NCA})(1 - (1 - \kappa_e) \lambda)\nu \tilde{V} - 1^{NCA}\kappa_c \nu \tilde{V} \right) \right\}.
\]

Let \( \bar{\kappa}_c(\kappa_e, \lambda) = 1 - (1 - \kappa_e)\lambda \). If \( z > 0 \), the incumbent maximizes her flow payoff by choosing \( 1^{NCA} \in \{0, 1\} \) which maximizes the term multiplying \( z \). Therefore, \( 1^{NCA} = 1 \) is strictly preferred iff \( 1 - (1 - \kappa_e)\lambda > \kappa_c \), which is equation (29).

### C.1.3 Proof of Proposition 4

**Proof.** First, suppose that Assumption 2 holds. To show existence, guess and verify that \( V(j, t|q) = \tilde{V} q \) and solve for equilibrium variables as described in the last section. Assumption 1 guarantees that household utility is finite and the obtained allocation is in fact an equilibrium. To show uniqueness, first notice that in any symmetric BGP, one has \( V(j, t|\bar{q}_{jt}) = \tilde{V} \bar{q}_{jt} \) by Proposition 2 and its corollary. Given this representation and \( \kappa_c \neq \bar{\kappa}_c \), Proposition 3 implies that all symmetric BGPs
have $1_{j}^{\text{NCA}} = 1^{\text{NCA}}$ for the same value for $1^{\text{NCA}}$. Using these facts in the equilibrium conditions above one can recursively compute all equilibrium variables.

Now suppose Assumption 2 does not hold. Then $z > 0$ contradicts equilibrium. A symmetric BGP therefore must have $z = 0$. Uniqueness is immediate, since we already know that $\tilde{z} = \tilde{L}_{RD}$ and $\tau^{S} = 0$ so we can compute all equilibrium variables uniquely following the steps in the main text. It remains to confirm that $z = 0$ is in fact optimal for the incumbent. First suppose that the first inequality fails, i.e. the usual steps imply $\tilde{z} < 0$. Because $\hat{w}_{RD} > 0$ in any equilibrium, the denominator $\chi(\lambda - 1) - (1 - 1^{\text{NCA}})(1 - (1 - \kappa_{e})\lambda)\nu - 1^{\text{NCA}}\kappa_{c}\nu$ is negative. Inspecting the incumbent’s FOC, this condition implies that there is no positive value $\tilde{V}$ such that $z > 0$ satisfies the incumbent’s FOC. Therefore, $z = 0$ is consistent with equilibrium. Next, suppose that the usual steps imply $\hat{z} > \bar{L}_{RD}$. Define $\xi$ as the value such that $\hat{w}_{RD}/\tilde{V} = \xi$ implies incumbents are indifferent about R&D. The fact that the expression for $\hat{z}$ implies $\hat{z} > \bar{L}_{RD}$ means that, in order for $\hat{z} = \bar{L}_{RD}$ to be consistent with entrant optimality, the R&D wage to incumbent value ratio must be higher. That is, in equilibrium $\hat{w}_{RD}/\tilde{V} > \xi$. This makes innovation strictly less profitable for the incumbent. Since she would have been indifferent when $\hat{w}_{RD}/\tilde{V} = \xi$, she strictly prefers $z = 0$ when $\hat{w}_{RD}/\tilde{V} > \xi$. This confirms that the constructed equilibrium is in fact an equilibrium. Finally, NCA policy is not determined by optimality because $z = 0$ implies that it is irrelevant to the incumbent’s optimization problem; and it has no effect on other equilibrium variables as its only effect is through spinout formation or the noncompete cost, both of which are zero when $z = 0$.

C.2 Additional derivations

C.2.1 Growth accounting equation

Let $\Delta > 0$ and let $J_{0}(\Delta)$ ($J_{1}(\Delta)$) denote the indices $j \in [0, 1]$ on which innovation occurs zero (one) times between $t$ and $t + \Delta$. By a law of large numbers applied to the continuum of random processes $\{\bar{q}_{jt}\}_{j \in [0, 1]}$ (see Uhlig, 1996 for more details), the set $J_{0}(\Delta)$ has measure $\mu_{1}\Delta = (\tau + \tau^{S} + \hat{\tau})\Delta + o(\Delta)$. The set $J_{0}(\Delta)$ has measure $1 - \mu_{1}\Delta + o(\Delta)$.

\[
\begin{align*}
Q_{t+\Delta} &= \int_{0}^{1} \bar{q}_{j,t+\Delta} dj = \int_{j \in J_{0}} \bar{q}_{jt} dj + \int_{j \in J_{1}} \lambda \bar{q}_{jt} dj + o(\Delta) \\
&= (1 - \mu_{1}\Delta - o(\Delta))Q_{t} + (\mu_{1}\Delta + o(\Delta))\lambda Q_{t} + o(\Delta) \\
&= (1 - \mu_{1}\Delta)Q_{t} + \mu_{1}\Delta \lambda Q_{t} + o(\Delta),
\end{align*}
\]  

where I used the fact that $\mathbb{E}[\bar{q}_{jt} | j \in J_{0}, t] = \mathbb{E}[\bar{q}_{jt} | j \in J_{1}, t] = Q_{t}$, since innovations happen at the same rate regardless of $\bar{q}_{jt}$. It follows that

\[
\frac{\dot{Q}_{t}}{Q_{t}} = \lim_{\Delta \to 0} \frac{Q_{t+\Delta} - Q_{t}}{\Delta} = (\lambda - 1)\mu_{1}.
\]
C.2.2 Multiplicity of equilibria

On the knife edge $\kappa_c = \bar{\kappa}_c$, there are multiple equilibria with distinct growth and welfare implications. This results from the fact that a positive mass of incumbents are indifferent about their NCA policy. The first proposition states that there are two symmetric BGPs where all incumbents choose the same policy.

Proposition C.1. If Assumptions 1 and 2 hold and $\kappa_c = \bar{\kappa}_c$ (as defined in Proposition 3), then:

1. There exist exactly two symmetric BGPs with $1_{jt}^{NCA} = 1_{jt}^{NCA}$: one with $1_{jt}^{NCA} = 0$ and one with $1_{jt}^{NCA} = 1$.

2. Both such equilibria have the same R&D labor allocations $z, \hat{z}$

3. The equilibrium with $1_{jt}^{NCA} = 0$ has a higher growth rate $g$

Proof. The proof of the first part is essentially the same as that of the previous proposition. The only difference is that either choice $1_{jt}^{NCA} = 1$ or $1_{jt}^{NCA} = 0$ is valid under Proposition 3. Given the representation $V(j,t|q) = \tilde{V}_q$ and the scaling of wages $\hat{w}_{RD,t} = \hat{w}_{RD}Q_t$ and $w_{RD,j}(1_{jt}^{NCA}) = w_{RD}(1_{jt}^{NCA})Q_t$, the derivation above uniquely determines uniquely the rest of the equilibrium conditional on $x$. This equilibrium has finite household utility as long as Assumption 1 holds. The second part follows from the fact that when $\kappa_c = \bar{\kappa}_c$, the expressions for equilibrium R&D effort $\hat{z}, \hat{z}$ do not depend on $1_{jt}^{NCA}$. The reason is that $1_{jt}^{NCA}$ only affects $\hat{z}, z$ through its effect on the incumbent’s effective wage, but here is the incumbent is indifferent between $1_{jt}^{NCA} = 1$ and $1_{jt}^{NCA} = 0$ hence faces the same effective wage. Mathematically, (33) has the expression $(1 - 1_{jt}^{NCA})(1 - (1 - \kappa_e)\lambda)\nu - 1_{jt}^{NCA}\kappa_e\nu = (1 - 1_{jt}^{NCA})\kappa_e\nu + 1_{jt}^{NCA}\kappa_e\nu$ in the denominator. Since $\kappa_c = \bar{\kappa}_c$, $\hat{z}$ is unaffected by $1_{jt}^{NCA}$, which in turn implies $z$ is also unaffected.

The last statement follows from the fact that $z, \hat{z}$ are the same in both equilibria, but $\tau^S = 0$ when $1_{jt}^{NCA} = 1$ and $\tau^S = \nu z I > 0$ when $1_{jt}^{NCA} = 0$. By the growth accounting equation (35), this implies $g$ is higher when $1_{jt}^{NCA} = 0$. 

The second proposition shows that there is a continuum of symmetric BGPs where a constant fraction of incumbents use NCAs.

Proposition C.2. If $0 \geq 1$, $\kappa_c = \bar{\kappa}_c$, and $\left(\frac{\lambda(1 - \kappa_c)\lambda}{(1 - \kappa_c)\lambda - \kappa_e\nu}\right)^{1/\psi} < \bar{L}_{RD}$, then for all $f \in (0,1)$ there exists a symmetric BGP in which, at any given time $t$, a fraction $f$ of incumbents $j$ have $1_{jt}^{NCA} = 1$.

Proof. Consider the generalized growth accounting equation (it simplifies to the one in the main text when $1_{jt}^{NCA} = 1_{jt}^{NCA}$),

$$g_t = (\lambda - 1)\left(\tau + \tilde{\tau} + z\nu \int_{j:1_{jt}^{NCA} = 0} \frac{\tilde{q}_{jt}dj}{Q_t}\right).$$

(74)

Unless the integral term is constant, then $g_t$ is non-constant, even with constant $\tilde{z}_{jt}, \hat{z}_{jt}$. The integral is equal to the product of the mass $m_t^0$ of goods whose incumbents choose $1_{jt}^{NCA} = 0$ and the average
relative quality of those goods \( \gamma_t^0 = E[\tilde{q}_t^0 | j^{NCA}_t = 0] \). Relative to the baseline model, the only substantial modification is that one needs to derive an expression for the evolution over time of

\[
\Gamma_t^x = m_t^x \gamma_t^x Q_t, \quad x \in \{0, 1\},
\]

and show that

\[
\frac{\dot{\Gamma}_t^0}{\Gamma_t^0} = \frac{\dot{\Gamma}_t^1}{\Gamma_t^1}.
\]

In this proof I will show that the integral remains constant as long as the random process \( j^{NCA}_t \) follows a certain stationary Markov process with states \( \{0, 1\} \). One such Markov process for \( j^{NCA}_t \) is to assume that \( j^{NCA}_t \) resets every time there is a new incumbent. To take the simplest scenario for the sake of exposition, suppose that the transition probabilities do not depend on the state. Specifically, suppose that, conditional on a transition, the likelihood of transitioning to \( j^{NCA} = 1 \) is \( p \in (0, 1) \). The equilibrium quantities \( \Gamma^0, \Gamma^1 \) evolve according to

\[
\Gamma^0_{t+\Delta} = \frac{\text{No innovations}}{(1 - \hat{\tau} + \nu z) \Delta \Gamma_t^0} + \frac{\text{Innovating incumbents}}{\Delta (\lambda - 1) \Gamma_t^0} + (1 - p) \lambda \left( \frac{\hat{\tau} + \nu z}{\Gamma_t^{NCA=0}} \right) + o(\Delta),
\]

\[
\Gamma^1_{t+\Delta} = \frac{\text{No innovations}}{(1 - \hat{\tau} + \nu z) \Delta \Gamma_t^1} + \frac{\text{Innovating incumbents}}{\Delta (\lambda - 1) \Gamma_t^1} + p \lambda \left( \frac{\hat{\tau} + \nu z}{\Gamma_t^{NCA=0}} \right) + o(\Delta),
\]

where \( o(\Delta) \) has the usual meaning that \( \lim_{\Delta \to 0} \frac{o(\Delta)}{\Delta} = 0 \). Subtracting \( \Gamma^x_t \), dividing by \( \Delta \), and taking the limit as \( \Delta \to 0 \) yields

\[
\dot{\Gamma}_t^0 = - (\hat{\tau} + \nu z) \Gamma_t^0 + \tau (\lambda - 1) \Gamma_t^0 + (1 - p) \lambda \left( \frac{\hat{\tau} + \nu z}{\Gamma_t^{NCA=0}} \right) \Gamma_t^1,
\]

\[
\dot{\Gamma}_t^1 = - (\hat{\tau} + \nu z) \Gamma_t^1 + \frac{p \lambda \left( \frac{\hat{\tau} + \nu z}{\Gamma_t^{NCA=0}} \right) \Gamma_t^0}{\Gamma_t^{NCA=1}}.
\]

Dividing by \( \Gamma_t^x \) yields

\[
\frac{\dot{\Gamma}_t^0}{\Gamma_t^0} = - (\hat{\tau} + \nu z) + \tau (\lambda - 1) + (1 - p) \lambda \left( \frac{\hat{\tau} + \nu z}{\Gamma_t^{NCA=0}} \right) \frac{\Gamma_t^1}{\Gamma_t^0},
\]

\[
\frac{\dot{\Gamma}_t^1}{\Gamma_t^1} = - (\hat{\tau} + \nu z) + \frac{p \lambda \left( \frac{\hat{\tau} + \nu z}{\Gamma_t^{NCA=0}} \right) \left( \frac{\Gamma_t^1}{\Gamma_t^0} \right)}{\Gamma_t^{NCA=1}}.
\]

Setting \( \dot{\Gamma}_t^0 = \dot{\Gamma}_t^1 \) and multiplying both sides by \( \frac{\Gamma_t^1}{\Gamma_t^0} \) yields a quadratic equation in \( \frac{\Gamma_t^1}{\Gamma_t^0} \), given by

\[
0 = \frac{a}{(1 - p) \lambda \hat{\tau}} \left( \frac{\Gamma_t^1}{\Gamma_t^0} \right)^2 + \frac{b}{(1 - p) \lambda (\hat{\tau} + \nu z)} \left( \frac{\Gamma_t^1}{\Gamma_t^0} \right) - \frac{c}{p \lambda (\hat{\tau} + \nu z)}.
\]
Using \( p \in (0, 1) \) and the facts that \( \lambda, \hat{\tau} > 0 \) and \( \nu, z \geq 0 \) yields \(-4ac > 0\). Then \( \frac{\Gamma^1}{\Gamma^0} = \frac{-b \pm \sqrt{b^2 - 4ac}}{2a} \) implies that there is always exactly one strictly positive real solution for \( \frac{\Gamma^0}{\Gamma^1} \). To see this, note that of course \( b^2 - 4ac > 0 \) so the all solutions are real. Then, \(-4ac > 0\) implies \( \sqrt{b^2 - 4ac} > |b| \). Regardless of whether \( b \) is positive or negative, \(-b \pm \sqrt{b^2 - 4ac} > 0\) and \( b \pm \sqrt{b^2 - 4ac} < 0\). The positive root is the equilibrium value of \( \Gamma^1_t \Gamma^0_t \). Using

\[
\Gamma^1_t + \Gamma^0_t = \int_{j:1}^{NCA} q_j t dj + \int_{j:1}^{NCA} \bar{q}_j t dj \\
= \int_0^1 q_j t dj \\
= Q_t,
\]

one has a linear system of two equations in two unknowns, \( \Gamma^1_t \) and \( \Gamma^0_t \). Given

\[
\frac{\Gamma^0_t}{Q_t} = \int_{j:1}^{NCA} q_j t dj,
\]

the BGP growth rate is given by (74) above.

C.3 Policy analysis derivations

C.3.1 R&D subsidy (tax)

Suppose that the planner subsidizes R&D spending at rate \( T_{RD} \) (tax if \( T_{RD} < 0 \)). By the same argument as before, Proposition 2 holds, so \( V(j, t|q) = \tilde{V} q \). Similarly, Lemma 4 holds. Therefore, the incumbent value \( \tilde{V} \) satisfies

\[
(r + \hat{\tau}) \tilde{V} = \hat{\pi} + \max_{z \geq 0} z \left\{ \chi(\lambda - 1) \tilde{V} - (1 - T_{RD})(\hat{\omega}_{RD} - (1 - 1^{NCA})(1 - \kappa_e) \nu \tilde{V}) \right\}.
\]

Then if \( z > 0 \), the incumbent’s optimal NCA policy is given by

\[
x = \begin{cases} 
1 & \text{if } \kappa_e < \tilde{\kappa}_e, \\
0 & \text{if } \kappa_e > \tilde{\kappa}_e, \\
\{0, 1\} & \text{if } \kappa_e = \tilde{\kappa}_e,
\end{cases}
\]

where \( \tilde{\kappa}_e = 1 - (1 - T_{RD})(1 - \kappa_e) \lambda \). Assuming \( z > 0 \), by the same logic as before one can obtain an expression for equilibrium \( \hat{z} \),

\[
\hat{z} = \left( \frac{\hat{\chi} (1 - \kappa_e) \lambda}{\chi(\lambda - 1) - \nu(1^{NCA} \kappa_e + (1 - 1^{NCA})(1 - (1 - T_{RD})(1 - \kappa_e) \lambda))} \right)^{1/\psi}.
\]
The rest of the equilibrium allocation and prices can be computed in the same way as before (including how to account for the possibility of $z = 0$), with the exception that the equilibrium R&D wage is now given by

$$\hat{w}_{RD} = (1 - T_{RD})^{-1}\hat{\chi}z^{-\psi}(1 - \kappa_e)\lambda \tilde{V}.$$ (88)

### C.3.2 Targeted R&D subsidy (tax)

Suppose that the planner subsidizes incumbent R&D spending at rate $T_{RD,I}$ (tax if $T_{RD,I} < 0$). By the same argument as before, Proposition 2 holds, so $V(j, t|q) = \tilde{V}q$. Similarly, Lemma 4 holds. Therefore, the incumbent’s value satisfies

$$ (r + \hat{\tau})\tilde{V} = \hat{\pi} + \max_{\{1, \}^{\hat{NCA} \in \{0, 1\}}} \left\{ z \left( \chi(\lambda - 1)\tilde{V} - (1 - T_{RD,I})\left( \hat{w}_{RD} - (1 - \hat{1}^{NCA})(1 - \kappa_e)\lambda \nu \tilde{V} \right) \right) \right\}.$$ (89)

It can be rearranged to a form analogous to (30),

$$ (r + \hat{\tau})\tilde{V} = \hat{\pi} + \max_{\{1, \}^{\hat{NCA} \in \{0, 1\}}} \left\{ \chi(\lambda - 1)\tilde{V} - (1 - T_{RD,I})\hat{w}_{RD} \right\}.$$ (90)

The noncompete policy is the same as before, with new threshold given by

$$\hat{\kappa}_c = 1 - (1 - T_{RD,I})(1 - \kappa_e).$$ (92)

If $z > 0$, using the same approach as before yields an expression for $\hat{z}$,

$$ \hat{z} = \left( \frac{(1 - T_{RD,I})\hat{\chi}(1 - \kappa_e)\lambda}{\chi(\lambda - 1) - \nu(\hat{1}^{NCA}\kappa_c + (1 - \hat{1}^{NCA})(1 - (1 - T_{RD,I})(1 - \kappa_e)\lambda))} \right)^{1/\psi}. $$ (93)
The remaining equilibrium conditions are

\[ \hat{\tau} = \hat{\chi} \hat{z}^{1-\psi}, \quad (94) \]
\[ z = \bar{L}_{RD} - \hat{z}, \quad (95) \]
\[ \tau = \chi z, \quad (96) \]
\[ \tau^S = (1 - 1^{NCA}) \nu z, \quad (97) \]
\[ g = (\lambda - 1)(\tau + \tau^S + \hat{\tau}), \quad (98) \]
\[ r = \theta g + \rho, \quad (99) \]
\[ \tilde{V} = \frac{\bar{\pi}}{r + \bar{\tau}}, \quad (100) \]
\[ \hat{w}_{RD} = \hat{\chi} \hat{z}^{-\psi}(1 - \kappa_e)\lambda \tilde{V}. \quad (101) \]

Finally, if \( \hat{z} \), as given by (93), violates \( \hat{z} \leq \bar{L}_{RD} \), then \( \hat{z} = \bar{L}_{RD} \) and \( z = 0 \). The rest of the equilibrium can be derived as is in the baseline case.

D Calibration appendix

D.1 Computing model moments

D.1.1 Profit to GDP

In the model, this ratio is simple to calculate using the solution to the static equilibrium as \( \bar{\pi}/Y \).

D.1.2 R&D to GDP

In the model, the R&D share is the ratio of the wage paid to R&D workers to GDP. This is

\[ \frac{\text{R&D wage bill}}{\text{GDP}} = \frac{w_{RD} z + \hat{w}_{RD} \hat{z}}{Y} = \frac{\hat{w}_{RD}(z + \hat{z}) + (w_{RD} - \hat{w}_{RD})z}{Y} = \frac{\hat{w}_{RD}(z + \hat{z}) - (1 - \kappa_e)\lambda \tilde{V} \tau^S}{Y}, \]

where I used \( w_{RD} - \hat{w}_{RD} = -(1 - 1^{NCA})(1 - \kappa_e)\lambda \tilde{V} \nu \) and \( \tau^S = (1 - 1^{NCA}) \nu z \).

D.1.3 Growth share OI

The model moment that corresponds here is the share of growth due to own innovation by incumbents of age \( \geq 6 \). In the model, the fraction of OI growth due to incumbents in a given age group is exactly their fraction of employment: innovations arrive at the same rate for each incumbent, and their impact on aggregate growth is proportional to the incumbent’s relative quality, which is proportional to employment. Hence old incumbents’ share of growth due to own innovation is simply
one minus the employment share calculated in the previous paragraph, $e^{((\tau_I-1)g-(\tilde{\tau}+(1-1^{NCA})z\nu))-6}$.

Finally, the fraction of aggregate growth due to OI is $\hat{\tau}_I$, defined above. The fraction of growth due to incumbents of age at least 6 is the product of the two,

$$\text{Age} \geq 6 \text{ share of OI} = \hat{\tau}_I \frac{\ell(6)}{\ell(0)} = \hat{\tau}_I e^{((\hat{\tau}_I-1)g-(\hat{\tau}+(1-1^{NCA})z\nu))-6}.$$  

D.1.4 Entry rate

Let $\ell(a)$ denote the density of incumbent employment at age $a$ incumbents. Then $\ell(a)$ is characterized by

$$\ell(a) = \ell(0)e^{((\hat{\tau}_I-1)g-(\hat{\tau}+\tau_S))a,}$$

$$1 - \hat{z} = \int_0^\infty \ell(a) da,$$

where $\hat{\tau}_I = \frac{\tau}{\tau+\hat{\tau}+\tau_S}$ is the fraction of innovations that are incumbents' own innovations.

The intuition for this characterization of $\ell(a)$ has two parts. First, because all shocks are iid across firms in equilibrium, the law of large numbers applied to each cohort of firms implies that we can consider directly the evolution of the cohort as a whole instead of explicitly analyzing the dynamics each individual firm in the cohort. Second, the employment of a firm is proportional to its relative quality, $l_j \propto q_j = q_j/Q$, as long as it is the leader. When it is no longer the leader, its employment is zero forever. Putting these two together, $\ell(a)$ must decline at exponential rate $g$ due to the increase in $Q_t$ (obsolescence), increase at rate $\hat{\tau}_I g$ due to incumbents own innovations, and decline at rate $\hat{\tau} + \tau_S$ due to creative destruction.\footnote{The second equation imposes consistency with aggregate employment; it implies $\ell(0) = -((\hat{\tau}_I-1)g-(\hat{\tau}+\tau_S))(1-\hat{z})$. The calibration does not require this explicit calculation since it is based only on employment shares.} Note that the employment density is strictly decreasing in $a$. This is because there are no adjustment costs: firms achieve their optimal scale immediately upon entry, and subsequently become obsolete (on average) or lose the innovation race to an entrant. Finally, due to the constant exponential decay of $\ell(a)$, the share of incumbent employment in incumbents of strictly less than 6 years of age is given by

$$\Xi_{[0,6]} = 1 - \frac{\ell(6)}{\ell(0)} = 1 - e^{((\hat{\tau}_I-1)g-(\hat{\tau}+\tau_S))6}.$$  

To calculate the share of employment in incumbents age $< 6$, I use as denominator the employment of intermediate goods firms in the economy (including R&D by entrants). When bringing this to the data, it is equivalent to assuming that the age-employment distribution of final goods firms is the same as that of intermediate goods firms. Using this approach, the share of overall employment in incumbents of age $< 6$, including R&D performed by non-producing entrants, is equal to the
previously calculated $\Xi_{[0,6]}$ multiplied by the share of total labor in incumbents $1 - L_F - \hat{z}$, added to the R&D labor used by entrants $\hat{z}$, divided by the share of total employment in intermediates $1 - L_F$, and finally multiplied by $2/3$ which is the share of creative destruction that corresponds to new firms in the data.\textsuperscript{58} This yields

$$\text{Age < 6 share of employment} = \frac{2}{3} \frac{\Xi_{[0,6]}(1 - L_F - \hat{z}) + \hat{z})}{1 - L_F}.$$ 

The factor $2/3$ deserves some additional discussion. According to Klenow and Li, 2020, creative destruction by incumbents is responsible for half as much growth as creative destruction by entrants. In this interpretation of the model, both types of creative destruction use the same technology. Therefore, it follows that $2/3$ of employment in young firms in the model represents employment in young firms in the data.

### D.1.5 Employment share of WSOs

Because successfully innovating spinouts and entrants have identical expected growth dynamics, the BGP share of employment in firms started as spinouts is their share of new incumbents $\tau^S \tau^S (\frac{2}{3}) \hat{z}$, multiplied by the employment share of incumbents $\frac{1 - L_F - (\frac{2}{3}) \hat{z}}{1 - L_F}$, and

$$\text{Spinout employment share} = \frac{\tau^S}{\tau^S + \frac{2}{3} \hat{z}} \times \frac{1 - L_F - (\frac{2}{3}) \hat{z}}{1 - L_F}.$$ 

Again, the factor $2/3$ reflects the fact that I assume $2/3$ of the entrants in the model are new firms in the data.

### E Policy analysis robustness appendix

#### E.1 NCA cost $\kappa_c$

#### E.1.1 Entry costs as transfers

Table E12 shows the effect on growth, the level of consumption, and welfare of setting $\kappa_c = 0$, treating the costs of entry as transfers to the competitive financial intermediary. The growth effect is the same as in Table 5, but the level of consumption is unaffected. This is because the aggregate cost of NCA enforcement is zero both when $\kappa_c = 0$ and when $\kappa_c > \bar{\kappa}_c$, as in the latter case NCAs are prohibitively expensive and not used in equilibrium. The result is that welfare increases due to the increased growth, albeit less as there are entry cost savings from reduced creative destruction.

\textsuperscript{58}Alternatively, one could assume that final goods firms have the same employment-age distribution as other intermediate goods firms. Then the formula would be

$$\text{Age < 6 share of employment} = \frac{2}{3} (\Xi_{[0,6]}(1 - \hat{z}) + \hat{z}).$$

This has only minor effects on the inferred parameters.
Table E12: Effect of reduction in $\kappa_c$ on growth, level of consumption, and welfare

<table>
<thead>
<tr>
<th>Measure</th>
<th>Variable</th>
<th>$\kappa_c &gt; \bar{\kappa}_c$</th>
<th>$\kappa_c = 0$</th>
<th>Chg.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Growth</td>
<td>$g$</td>
<td>1.487%</td>
<td>1.695%</td>
<td>0.21 p.p.</td>
</tr>
<tr>
<td>Level</td>
<td>$C$</td>
<td>0.80</td>
<td>0.80</td>
<td>0%</td>
</tr>
<tr>
<td>Welfare</td>
<td>$W$</td>
<td></td>
<td>2.86% (CE)</td>
<td></td>
</tr>
</tbody>
</table>

Figure E10 plots the determinants of the growth rate and level of consumption for $\kappa_c \in [0, 2\bar{\kappa}_c]$. For $0 < \kappa_c < \bar{\kappa}_c$, the aggregate direct cost of NCAs is positive. Still, the overall magnitude is quite small, reducing the level of consumption by at most 0.2 percentage points.

### E.1.2 Incumbent decreasing returns to scale in innovation

As discussed in Section 5.2.1, the magnitude of the increase in growth and welfare would be smaller if the price-elasticity of incumbent demand for R&D labor were lower. In the baseline model, this elasticity is infinite as the incumbent has constant returns to scale. In this section I consider an extended model which specifies the innovation production functions as

$$
\tau_{jt} = \chi z_{jt}^{1-\psi},
$$

$$
\tau_{jt}^S = (1 - 1^{NCA} j)\nu z_{jt},
$$

$$
\hat{\tau}_{jt} = \hat{\chi} z_{jt}^{1-\hat{\psi}},
$$

where I now use the notation $\hat{\psi}$ to refer to the parameter corresponding to $\psi$ in the baseline model, in order to have consistent notation (hats always refer to entrants). Note that I keep the rate of spinout formation linear in $z_{jt}$. This is necessary for tractability as otherwise the wage paid by the incumbent would depend on $z_{jt}$: specifically, a higher $z_{jt}$ would imply a lower value of future spinouts per worker and hence require a higher wage.

**Equilibrium conditions** The equilibrium conditions of this model are the same as before except for the incumbent HJB.\(^{59}\) In particular, note that the use of NCAs in equilibrium again follows (29) so it can be derived in closed form. In normalized terms (i.e., $V(j, t|\bar{\bar{q}}_{jt}) = \tilde{V} \tilde{q}_{jt}$), the incumbent HJB is

$$
(r + \hat{\tau})\tilde{V} = \tilde{\pi} + \max_{z \geq 0, 1^{NCA} \in \{0, 1\}} \left\{ \chi z^{1-\psi}(\lambda - 1)\tilde{V} - zw_{RD}(1^{NCA}) - (1 - 1^{NCA})z\nu \tilde{V} - 1^{NCA} z\kappa_c \nu \tilde{V} \right\}.
$$

\(^{59}\)I have not proven uniqueness of the symmetric BGP in this case; however, numerically there is no evidence so far of multiplicity.
Figure E10: Effect of varying $\kappa_c$ on key equilibrium variables, considering the entry cost as a transfer to the financial intermediary. The top-left panel shows R&D labor allocated to incumbents (own-product innovation) and entrants (creative destruction). The top-right panel shows the aggregate productivity growth rate. The bottom-left panel shows the direct cost of NCAs. Finally, the bottom-right panel shows the level of consumption.
The first order condition with respect to \( z \) is given by

\[
(1 - \psi) \chi z^{-\psi} (\lambda - 1) \hat{V} = w_{RD}(1^{NCA}) + (1 - 1^{NCA}) \nu \hat{V} - 1^{NCA} \kappa_c \nu \hat{V},
\]

which yields an expression for \( z(\hat{V}, w_{RD}(1^{NCA}), 1^{NCA}) \),

\[
z(\hat{V}, w_{RD}(1^{NCA}), 1^{NCA}) = \left( \frac{(1 - \psi) \chi (\lambda - 1) \hat{V}}{w_{RD}(1^{NCA}) + (1 - 1^{NCA}) \nu \hat{V} - 1^{NCA} \kappa_c \nu \hat{V}} \right)^{\frac{1}{\nu}}.
\]

Plugging (107) into the incumbent HJB yields an equilibrium relationship between \( \hat{V}, w_{RD}(1^{NCA}) \) (given \( r, \hat{\tau}, \hat{\pi}, 1^{NCA} \) in the background). Suppressing the dependence of \( w_{RD} \) on \( 1^{NCA} \) for clarity, this is

\[
(r + \hat{\tau}) \hat{V} = \hat{\pi} + \chi z(\hat{V}, w_{RD}, 1^{NCA})^{1-\psi} (\lambda - 1) \hat{V} - z(\hat{V}, w_{RD}, 1^{NCA})(w_{RD} + (1 - 1^{NCA}) \nu \hat{V} - 1^{NCA} \kappa_c \nu \hat{V}).
\]

Solution algorithm  
I use the following algorithm to compute a symmetric BGP.

1. Use (29) to determine \( 1^{NCA} \).
2. Guess \( r, \hat{\tau} \).
3. Use nonlinear root finding algorithm to compute \( \hat{V}, w_{RD}(1^{NCA}), \hat{w}_{RD}, z, \hat{\varepsilon} \) by simultaneously solving (108), entrant optimization, worker indifference condition for R&D labor supply, and R&D labor market clearing.
4. Compute \( \tau, \hat{\tau}, \tau^S \) and \( g = (\lambda - 1)(\tau + \hat{\tau} + \tau^S) \).
5. Compute \( r = \theta g + \rho \) (Euler equation).
6. Check whether \( r, \hat{\tau} \) are within tolerance of guesses from Step 2. If so, an equilibrium has been found. If not, update and return to Step 2.

Results  
To study the properties of this model, I recalibrate to the same target moments. I assume \( \psi = 0.5\) which is a standard choice in the literature. As discussed in Akcigit and Kerr, 2018, this value of \( \psi \) is consistent with the price-elasticity of R&D that has been found in the micro data. The calibration again is able to match the moments exactly. The inferred parameters, however, are different than in the model with a constant returns to scale production function. They are listed in Table E13.

Table E14 shows the results of reducing \( \kappa_c \) to zero in this model. Growth increases by 0.18 percentage points and welfare increases by 2.65% in CE terms. This is slightly weaker than the results in the case with constant returns to scale. Table E15 decomposes the sources of growth in the high and low \( \kappa_c \) BGPs. As in the baseline model, \( \kappa_c = 0 \) has a higher allocation of R&D to own-product innovation. This increase growth for the same reason as in the baseline case. Figure E11 traces these effects for all values of \( \kappa_c \) between \( \bar{\kappa}_c \) and zero.

82
Table E13: Calibrated parameters with DRS incumbent innovation

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Value</th>
<th>Description</th>
<th>Source</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\theta$</td>
<td>2</td>
<td>$\theta^{-1} = \text{IES}$</td>
<td>External</td>
</tr>
<tr>
<td>$\psi$</td>
<td>0.5</td>
<td>Incumbent R&amp;D curvature</td>
<td>External</td>
</tr>
<tr>
<td>$\hat{\psi}$</td>
<td>0.5</td>
<td>Entrant R&amp;D curvature</td>
<td>External</td>
</tr>
<tr>
<td>$\rho$</td>
<td>0.0559</td>
<td>Discount rate</td>
<td>Internal</td>
</tr>
<tr>
<td>$\beta$</td>
<td>0.094</td>
<td>$\beta^{-1} = \text{EoS intermediate goods}$</td>
<td>Internal</td>
</tr>
<tr>
<td>$\lambda$</td>
<td>1.087</td>
<td>Quality ladder step size</td>
<td>Internal</td>
</tr>
<tr>
<td>$\chi$</td>
<td>2.73</td>
<td>Incumbent R&amp;D productivity</td>
<td>Internal</td>
</tr>
<tr>
<td>$\hat{\chi}$</td>
<td>0.356</td>
<td>Entrant R&amp;D productivity</td>
<td>Internal</td>
</tr>
<tr>
<td>$\kappa_e$</td>
<td>0.587</td>
<td>Non-R&amp;D entry cost</td>
<td>Internal</td>
</tr>
<tr>
<td>$\nu$</td>
<td>0.900</td>
<td>Spinout generation rate</td>
<td>Internal</td>
</tr>
<tr>
<td>$\bar{L}_{RD}$</td>
<td>0.01</td>
<td>R&amp;D labor allocation</td>
<td>Internal</td>
</tr>
</tbody>
</table>

Table E14: Effect of reducing $\kappa_c$ in DRS incumbent innovation calibration

<table>
<thead>
<tr>
<th>Measure</th>
<th>Variable</th>
<th>$\kappa_c &gt; \bar{k}_c$</th>
<th>$\kappa_c = 0$</th>
<th>Chg.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Growth</td>
<td>$g$</td>
<td>1.487%</td>
<td>1.664%</td>
<td>0.177 p.p.</td>
</tr>
<tr>
<td>Level</td>
<td>$\bar{C}$</td>
<td>0.786</td>
<td>0.788</td>
<td>0.25%</td>
</tr>
<tr>
<td>Welfare</td>
<td>$\bar{W}$</td>
<td></td>
<td></td>
<td>2.65% (CE)</td>
</tr>
</tbody>
</table>

Table E15: Decomposition of effect of reducing $\kappa_c$ on growth and R&D in DRS incumbent innovation calibration

<table>
<thead>
<tr>
<th>Measure</th>
<th>Variable</th>
<th>$\kappa_c &gt; \bar{k}_c$</th>
<th>$\kappa_c = 0$</th>
<th>Chg.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Growth</td>
<td>$g$</td>
<td>1.487%</td>
<td>1.664%</td>
<td>0.177 p.p.</td>
</tr>
<tr>
<td></td>
<td>$(\lambda - 1)\tau$</td>
<td>1.20%</td>
<td>1.41%</td>
<td>0.21 p.p.</td>
</tr>
<tr>
<td></td>
<td>$(\lambda - 1)\hat{\tau}$</td>
<td>0.268%</td>
<td>0.249%</td>
<td>-0.019 p.p.</td>
</tr>
<tr>
<td></td>
<td>$(\lambda - 1)\tau^S$</td>
<td>0.020%</td>
<td>0%</td>
<td>-0.02 p.p.</td>
</tr>
<tr>
<td>R&amp;D</td>
<td>incumbents (%)</td>
<td>$z/L_{RD}$</td>
<td>25.51%</td>
<td>35.49%</td>
</tr>
<tr>
<td></td>
<td>entrants (%)</td>
<td>$\hat{z}/L_{RD}$</td>
<td>74.49%</td>
<td>64.51%</td>
</tr>
</tbody>
</table>
Figure E11: Effect of varying $\kappa_c$ on key equilibrium variables in the model with incumbent decreasing returns to scale in innovation. The top-left panel shows R&D labor allocated to incumbents (own-product innovation) and entrants (creative destruction). The top-right panel shows the aggregate productivity growth rate. The bottom-left panel shows the direct cost of NCAs. The bottom-right panel shows the level of consumption.
\[ \nabla_{\log m} (CE \text{ Welfare chg.\%}) \]

Figure E12: Sensitivity of welfare comparison to moments. This is \((J^{-1})^T \nabla_p W\), where \(W(p)\) maps log parameters to the log of the percentage change in BGP consumption which is equivalent to the change in welfare from changing \(\kappa_c\) from \(\infty\) to 0 (i.e. going from banning to frictionlessly enforcing NCAs).

### E.1.3 Variation in target moments

Figure E12 shows the sensitivity of the welfare comparison the moments targeted, including the externally calibrated parameters as pseudo-moments as before. As discussed in the main text, it is computed as \(\nabla_{\log m} \tilde{W}|_{\log m} = (J^{-1})^T \nabla_{\log p} W|_{\log p}\), where \(J\) is the Jacobian of the mapping from log parameters to moments (so that \(J^{-1}\) is the Jacobian of the inverse mapping), and \(W\) is the mapping from log parameters the consumption-equivalent percent change in welfare from reducing \(\kappa_c\) from \(\kappa_c > \bar{\kappa}_c\) to \(\kappa_c = 0\). That is, it is the gradient of the change in welfare to the log change in target moments or uncalibrated parameters, taking as given the change in parameters required to continue matching the target moments. For reference, \(\nabla_p W|_p\) for each definition of \(W\) can be found in Figure B6.

To get a sense of what this means about robustness of the results, suppose that the log of each moment is assumed to have a standard deviation of \(\sigma\) and that this uncertainty is uncorrelated across moments. The uncertainty propagates such that the standard deviation of the CE welfare change is the square root of \((\nabla_m \tilde{W}|_m)^T \Sigma_m \nabla_m \tilde{W}|_m\), where \(\Sigma_m = \sigma^2 I_{9 \times 9}\). In this examples this yields 13.2\(\sigma\) standard deviation of the welfare improvement, measured in percentage points. Given a baseline welfare improvement of 3.24\%, a standard deviation excludes zero for \(\sigma \leq .245\) or about 21.5\%. Two standard deviations exclude zero for \(\sigma \leq .123\) or about 11.6\%.
Table E16: Low entry rate calibration

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Value</th>
<th>Description</th>
<th>Source</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\theta$</td>
<td>2</td>
<td>$\theta^{-1} =$ IES</td>
<td>External</td>
</tr>
<tr>
<td>$\psi$</td>
<td>0.5</td>
<td>Entrant R&amp;D curvature</td>
<td>External</td>
</tr>
<tr>
<td>$\rho$</td>
<td>0.0559</td>
<td>Discount rate</td>
<td>Internal</td>
</tr>
<tr>
<td>$\beta$</td>
<td>0.094</td>
<td>$\beta^{-1} =$ EoS intermediate goods</td>
<td>Internal</td>
</tr>
<tr>
<td>$\lambda$</td>
<td>1.56</td>
<td>Quality ladder step size</td>
<td>Internal</td>
</tr>
<tr>
<td>$\chi$</td>
<td>2.77</td>
<td>Incumbent R&amp;D productivity</td>
<td>Internal</td>
</tr>
<tr>
<td>$\check{\chi}$</td>
<td>0.131</td>
<td>Entrant R&amp;D productivity</td>
<td>Internal</td>
</tr>
<tr>
<td>$\kappa_e$</td>
<td>0.577</td>
<td>Non-R&amp;D entry cost</td>
<td>Internal</td>
</tr>
<tr>
<td>$\nu$</td>
<td>0.082</td>
<td>Spinout generation ral</td>
<td>Internal</td>
</tr>
<tr>
<td>$\bar{L}_{RD}$</td>
<td>0.01</td>
<td>R&amp;D labor allocation</td>
<td>Internal</td>
</tr>
</tbody>
</table>

Table E17: Effect of reducing $\kappa_c$ in low entry calibration

<table>
<thead>
<tr>
<th>Measure</th>
<th>Variable</th>
<th>$\kappa_c &gt; \bar{\kappa}_c$</th>
<th>$\kappa_c = 0$</th>
<th>Chg.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Growth</td>
<td>$g$</td>
<td>1.487%</td>
<td>1.466%</td>
<td>−0.021 p.p.</td>
</tr>
<tr>
<td>Level</td>
<td>$\bar{C}$</td>
<td>0.793</td>
<td>0.794</td>
<td>0.13%</td>
</tr>
<tr>
<td>Welfare</td>
<td>$\bar{W}$</td>
<td></td>
<td>−0.23% (CE)</td>
<td></td>
</tr>
</tbody>
</table>

E.1.4 When is $\kappa_c = 0$ bad for welfare?

The sensitivity of the welfare improvement to the entry rate shown in Figure E12 suggests that a calibration targeting a lower entry rate – but, crucially, with the same share of growth coming from young firms – could have the opposite result. Figure E13 shows the analogue of Figure B7 if entry rate targeted is 5% instead of 13.34%. The model is again able to match the moments exactly; inferred parameter values are shown in Table E16. In this low entry calibration, growth and welfare fall when $\kappa_C$ is reduced to zero. The lower employment in young firms (while holding constant the fraction of growth coming from old firms) means that each entry by a young firm must have a higher effect on growth in order for the model to match the growth rate. Furthermore, as shown in Figure B6, the increase in $\lambda$ eliminates the overall welfare gain from reducing $\kappa_c$. As discussed previously, a higher value of $\lambda$ brings (54) closer to unity, weakening the growth increase from reallocation of R&D to own-product innovation.
Figure E13: Effect of varying $\kappa_c$ on equilibrium objects when the model is calibrated to a 5% share of employment in young firms instead of 13.34% as in the baseline calibration.
Table E18: Decomposition of effect of reducing $\kappa_c$ on growth and R&D in low entry calibration

<table>
<thead>
<tr>
<th>Measure</th>
<th>Variable</th>
<th>$\kappa_c &gt; \tilde{\kappa}_c$</th>
<th>$\kappa_c = 0$</th>
<th>Chg.</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Growth</strong></td>
<td>$g$</td>
<td>1.487%</td>
<td>1.466%</td>
<td>−0.021 p.p.</td>
</tr>
<tr>
<td>incumbents</td>
<td>$(\lambda - 1)\tau$</td>
<td>1.04%</td>
<td>1.06%</td>
<td>0.02 p.p.</td>
</tr>
<tr>
<td>entrants</td>
<td>$(\lambda - 1)\tilde{\tau}$</td>
<td>0.415%</td>
<td>0.407%</td>
<td>−0.008 p.p.</td>
</tr>
<tr>
<td>spinouts</td>
<td>$(\lambda - 1)\tau^S$</td>
<td>0.03%</td>
<td>0%</td>
<td>−0.03 p.p.</td>
</tr>
<tr>
<td><strong>R&amp;D</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>incumbents (%)</td>
<td>$z/\bar{L}_{RD}$</td>
<td>67.6%</td>
<td>68.8%</td>
<td>1.2 p.p.</td>
</tr>
<tr>
<td>entrants (%)</td>
<td>$\hat{z}/\bar{L}_{RD}$</td>
<td>32.4%</td>
<td>31.2%</td>
<td>−1.2 p.p.</td>
</tr>
</tbody>
</table>

E.2 NCA cost $\kappa_c$ and targeted R&D subsidy

E.2.1 Entry costs as transfers

Here, I again consider how the result is changed when entry costs are interpreted as transfers to the financial intermediary. The analogue of Figure 7 is shown in Figure E14. The optimal welfare improvement is now approximately 10.1%, which is smaller as there is no longer a welfare improvement from a reduction in a resource cost associated with creative destruction. For targeted R&D subsidies above 62%, it is optimal to ban NCAs, compared to a threshold of 77% subsidies in the baseline interpretation of entry costs. The reason is that creative destruction is relatively less inefficient as some of its costs are transfers rather than scarce resource costs. This means that a smaller reallocation of R&D achieves the social optimum and that sacrificing entry by spinouts entails a larger social cost.
Figure E14: Summary of equilibrium for baseline parameter values and various values of $T_{RD,I}$ and $\kappa_c$, considering entry costs as transfers.
References


Agarwal, Rajshree et al. (Feb. 1, 2002). “Knowledge Transfer through Congenital Learning: Spin-Out Generation, Growth and Survival”.


