January 2019

Synergizing Ventures*

by

Ufuk Akcigit, Emin Dinlersoz, Jeremy Greenwood, and Veronika Penciakova

Chicago, U.S. Census Bureau, Penn, and Maryland

Abstract

Venture capital and growth are examined both empirically and theoretically. Empirically, VC-backed startups have higher early growth rates and patenting levels than non-VC-backed ones. Venture capitalists increase a startup's likelihood of reaching the right tails of firm size and innovation distributions. Furthermore, there is positive assortative matching: better venture capitalists match with better startups, creating a synergistic effect. An endogenous growth model, where venture capitalists provide both expertise and financing to business startups, is constructed to match these facts. The degree of assortative matching and the taxation of VC-backed startups are important for growth.

Keywords: endogenous growth, IPO, management, mergers and acquisitions, research and development, selection effects, startups, assortative matching, synergies, taxation, treatment effects, venture capital

---

* *Work in progress.* Any opinions and conclusions expressed herein are those of the authors and do not necessarily represent the views of the U.S. Census Bureau. All results have been reviewed to ensure that no confidential information is disclosed.

Figure 1:

1 Introduction

While most firms are born small and stay small throughout their lifetimes, a few become large, successful businesses that not only dominate their industries, but also generate large externalities for the rest of the economy through innovations they create, and the extensive input-output networks they form. Furthermore, the fluctuations these large firms experience matter for the aggregate economy.[1] Venture capital (VC) is one of the catalysts that can take startups to the right tail of the firm-size and innovation distributions, and turn them into engines of growth. How much role does VC play in the making of these large and productive firms? How critical is VC funding and nurturing for growth and innovation rates of startups and the economy as a whole? To what extent do the synergies between entrepreneurs and venture capitalists matter for firm and aggregate growth? Answers are provided to these questions using both empirical analysis and a model of innovation and growth where venture capitalists match with startups. Policies that enhance the degree of assortative match between venture capitalists and startups, and those that allow the taxation of VC-backed firms at a lower rate than others, increase growth significantly.

Certain key events in the life cycle of firms appear to be precursors of high growth. There is a growing interest in identifying these precursors. For instance, having a patent or a trademark application early on is strongly predictive of the ultimate success of startups, as measured by rare events such as an IPO or a high-value acquisition—on this, see Fazio et

[1] These firms can also account for a large portion of fluctuations and growth in the U.S. economy. Gabaix (2011) analyzes this "granular hypothesis" and finds that the shocks experienced by the top 100 firms in the "fat" right tail of the firm-size distribution account for a sizable fraction of year-to-year GDP fluctuations.

al. (2016). Similarly, patenting and R&D activity are positively correlated with measures of firm growth—e.g., Foster, Grim and Zolas (2016). Furthermore, initial firm size is a good predictor of subsequent growth and success of a firm (Brown et al. (2017)), and ex ante differences among startups can account for a large amount of ex post firm size heterogeneity at any firm age (Pugsley, Sedlacek, and Sterk (2017)). Some early characteristics of planned businesses also predict well the ultimate formation of a business and its initial size—see Bayard et al. (2018). VC-backing also appears to be a key precursor of startup success. As shown below, there are both large selection and treatment effects on growth and innovation associated with VC presence in a startup—VC targets more promising startups very early in their life-cycles, and contributes significantly to their ultimate performance.

The dominance of large and productive firms ("superstar firms") in the U.S. economy has been increasing—as documented by Autor et al. (2017a, 2017b). The consequences include the reallocation of labor to superstar firms, accompanied by a rising industry concentration of sales and a fall in labor's share of income. The mechanisms by which superstar firms emerge, however, have not been explored in detail. In particular, the role of VC in taking startups to stardom by enhancing their management and financing their research and development has not been quantified. On the one hand, venture capitalists may invest in many risky projects with a high likelihood of failure. On the other hand, the growth rate of few surviving VC-backed projects may be exceptionally high, justifying spreading investment over several risky projects. VC funding and expertise can enhance the growth and innovativeness of a startup to a level that not only leads to a large and influential business, but also generates higher growth and innovation in the economy as a whole through the innovation externalities such firms create. The findings here suggest that VC-backed firm are indeed much more likely to eventually land in the right tail of firm size and innovation distributions, and become large and innovative firms.

The analysis here combines data on firms receiving VC funding with Census micro data for the period 1980-2012. The goal is to study the life-cycle dynamics of VC-funded versus non-VC-funded firms and understand the differences between them in terms of innovativeness and growth. Towards that goal, several new analyses are performed. First, at the heart of the model studied here is the synergies between venture capitalists and entrepreneurs in enhancing startup success and growth. Empirical analysis explores the role of assortative matching between venture capitalists and entrepreneurs in generating selection and treatment effects associated with VC. Startups that have more promising prospects tend to match with venture capitalists that have more expertise in funding projects and can offer more funding on average. Furthermore, firms backed by venture capitalists with higher funding capabilities and expertise tend to achieve significantly better outcomes.

Second, because innovation is a key ingredient of productivity growth in the model, firm-level data on patent applications and citations are used to study the effects of VC involvement on firm innovation. The model emphasizes the role of VC in fostering firm R&D and innovation. The number and quality of patents by VC-funded and non-VC-funded firms are compared to understand whether VC-funding disproportionately targets more innovative startups and also leads to higher innovation. In addition, the data on innovation is used to quantify the externalities imposed by VC-funded firms on the productivity and growth of the rest of the economy.

Finally, a model of VC funding is developed that can capture many of stylized facts presented. The model emphasizes the roles of VC funding of innovation and synergies between venture capitalists and entrepreneurs in generating firm success and economic growth. The model is used to assess the effects on growth of the degree of assortative matching between venture capitalists and entrepreneurs and the differential taxation of VC-backed startups versus others.

There is a large and growing literature on the effects of VC funding on firm performance. Puri and Zarutskie (2012) find that VC-backed firms exhibit higher growth and have much higher rates of IPO and M&A activity, compared to otherwise similar firms. Catalini, Guzman, and Stern (2017) estimate that startups that receive VC-funding are nearly 500 percent more likely to end up with an IPO or high-profile acquisition, controlling for observable characteristics that determine VC selection.[2] They also find that early firm-level predictors of

[2] Unlike Puri and Zarutskie (2012) and Catalini, Guzman and Stern (2017), who measure startup performance by traditional discrete outcomes (IPO and M&A), the analysis here uses, in addition, continuous measures of firm growth in employment, revenue, and patenting to define thresholds for success by the end of a typical VC contract. In particular, the likelihood of a venture capitalist taking a firm to the right tail of the firm-size and innovation distributions (90th percentile) within 10 years is assessed. The rationale

VC-funding are highly correlated with early predictors of ultimate firm success in the form of an IPO or acquisition, leaving a potentially very large role for VC selection. There is also evidence that VC involvement in a firm in the form of monitoring matters. Using an exogenous reduction in VC oversight costs due to the opening of new airline routes that make it easier for venture capitalists to visit their portfolio companies, Bernstein, Giroud, and Townsend (2015) show that venture capitalists' on-site involvement with the firms they fund generates both a higher level of innovation and a higher likelihood of success. These findings point to potentially large treatment effects associated with venture capitalists' monitoring and mentoring of startups.

The analysis here complements the prior empirical work summarized above, and extends it in a number of new dimensions. The focus here is on the specific roles of VC selection and treatment in startup innovation, productivity, and growth, and the effects of assortative matching between entrepreneurs and venture capitalists. The selection by venture capitalists based on early innovation and growth of startups is studied, in addition to the effects of VC treatment on innovation and growth. Innovative externalities generated by VC-backed firms are quantified using citations to the patents applied for by these firms. In addition, the presence and strength of assortative matching is assessed empirically. More importantly, a macroeconomic model is proposed to match some salient features of VC in the U.S. data. The model is calibrated where venture capitalists nurture talented entrepreneurs by providing the necessary ingredients, advice and money, to bring a startup to the market. The framework stresses the complementarity between a talented entrepreneur's skill and a venture capitalist's. Not all talented entrepreneurs can find a venture capitalist to back their startup. As a result, they turn to more traditional sources of financing, such as banks. Banks provide a limited amount of expertise for a startup. For a talented entrepreneur matching with a venture capitalist, as opposed to a bank, will result in higher probabilities of success, greater levels of funding for startup research and development, and higher levels of productivity. En-

for doing so is that even though a surviving VC-backed firm may not reach an IPO or M&A stage, it can still achieve substantial growth and innovation on average, compared to non-VC-backed firms. Furthermore, among the firms that are involved in IPO or M&A activity, the heterogeneity in performance can be more finely captured by continuous measures of firm growth. Measuring success in this fashion accounts for the contribution of all VC-backed firms to economic growth and aggregate productivity.

trepreneurs starting more mundane enterprises are more likely to use bank financing. The opportunity cost for a venture capitalist is too high to mentor this type of startup. The calibrated model is then used to assess how assortative matching between entrepreneurs and venture capitalists and the taxation of startups affect economic growth.

There is an emerging literature in macroeconomics on VC. Silveira and Wright (2016) build a canonical search model of the process where entrepreneurs are matched with venture capitalists. Upon meeting, the parties bargain in Nash fashion over each one's investment and how to split the proceeds. The analysis is theoretical in nature. Jovanovic and Szentes (2013) focus on a setting where the incubation period for a project is unknown. Unlike entrepreneurs, venture capitalists have deep pockets and can weather supporting a project over a prolonged period of time, if they so choose. A contract specifies the initial investment by the venture capitalist and some fixed split of the profits. The analysis focuses on characterizing and measuring the excess returns earned by venture capitalists, due to the latters' scarcity. Both of the above two papers abstract from innovation and growth.

Greenwood, Han, and Sanchez (2018) and Opp (2018) build growth models that incorporate VC. Opp (2018) develops a tractable stylized Schumpeterian model of VC that has analytical solutions. He estimates that the welfare benefits of VC are worth 1 to 2 percent of aggregate consumption, despite the fact that VC investment is highly procyclical, which operates to trim the estimates. In his analysis the research and development decisions by entrepreneurs are limited in the sense that they do not choose how far to launch their endeavor ahead of the pack. Also, the likelihood of success does not depend on the level of development funding. Greenwood, Han, and Sanchez (2018) model VC from a dynamic contract perspective. Their analysis focuses on modeling the VC funding round process. They don't explicitly model a banking sector as is done here. Neither Greenwood, Han, and Sanchez (2018) or Opp (2018) focuses on modelling the initial selection process between different types of entrepreneurs and financiers and the associated transition from ex ante promise to ex post success; that is, on the unique set of facts assembled here.

2 Some Stylized Facts on VC-Funded Firms

To what extent does VC matter in turning startups into engines of growth and innovation in the economy? Empirical analysis is carried out to offer some evidence on both the selection and treatment effects associated with venture capitalist involvement in a firm. This involvement may include, in addition to financial support, providing management advice and mentoring, industry expertise, and business network connections. These factors can lead to potentially different trajectories of growth and innovation for VC-funded firms, compared to others.

The first goal of the empirical analysis is to explore the nature of selection by venture capitalists, based on key observable characteristics of startups such as size, growth, and innovative activity early in their life-cycles. If venture capitalists tend to pick more promising projects to fund and nurture, there should be a disproportionate presence of VC among nascent projects that are initially larger, growing faster, and innovating more.

The second goal is to understand the role VC plays in catapulting firms to the right tail of firm size and innovation distributions. How much does VC matter for a startup's survival, growth, and innovative activity? To investigate such effects, the outcomes of VC-funded firms are compared with those of a control group composed of non-VC-funded firms, which are chosen to closely match the funded firms on a number of key characteristics at the time of VC funding.

Finally, the empirical analysis also investigates the degree to which assortative matching prevails among venture capitalists and startups. Some venture capitalists may be better than others at identifying and nurturing promising startups. That is, the extent of selection and treatment by venture capitalists may depend on *both* the type of startup *and* the type of venture capitalist. Are startups with better growth and innovation prospects disproportionately targeted by more able venture capitalists? Do startups funded and supported by more capable venture capitalists achieve better outcomes? Answers are provided by studying the outcomes of startups backed by high- versus low-quality venture capitalists, where quality is measured using various characteristics of venture capitalists.

2.1 Data

To carry out the analysis, a large number of datasets are brought together—see Appendix 10 for details on the construction of the analysis dataset, and all tables and figures. The main dataset that enables the tracking of firm outcomes over time is the U.S. Census Bureau's Longitudinal Business Database (LBD). The LBD contains longitudinally-linked information for each employer business in the U.S. starting in 1976 based on administrative data in the Census Bureau's Business Register (BR)—the main sampling frame for Census Bureau's surveys. Firms, and their constituent establishments, are observed in the LBD starting from the time of their first appearance as an employer business in administrative records. The LBD contains annual measures of employment, age, and payroll for both firms and establishments, and an annual revenue measure at the firm level for a large fraction of firms. (Data on revenue starts in 1997.)

Firms that receive VC funding are identified using the Thomson Reuters' VentureXpert database, which contains information on the dates of VC funding, the funding amount, and the identity of the venture capitalists. Firms in the VentureXpert dataset are first matched with the Census Bureau's Business Register (BR) using name and address matching techniques, and then linked to the LBD using unique firm identifiers. To include measures of innovation, the U.S. Census Bureau's patent-firm crosswalk is used to identify firms with patenting activity.[3] Citations for each patent are also added from USPTO's PatentsView. In addition, Standard&Poor's Compustat database is linked with the LBD to identify publicly-listed firms. The analysis sample based on the constructed dataset covers all employer firms born during the period 1980 to 2012. (Agricultural and government-sector firms are excluded from the analysis.)

A critical feature of the constructed dataset is that firms can be tracked over their entire life cycle, starting from their birth, as identified by the year of their first employment record, all the way to their death, which happens if the firm ceases to exist in the administrative data for employment. Other key events in a firm's life cycle are also observed: their first patent application, their first injection of funds by an venture capitalist, and an IPO or a

[3]The patent-firm crosswalk is a result of merging the USPTO's patent database with the LBD. We thank Nick Zolas for providing this crosswalk. See also Graham et al. (2015).

M&A. IPO events are identified using the Compustat database. M&A activity is identified using a set of criteria based on how the ownership of the firm and its establishments changes from one year to another.[4] The main firm outcomes in the analysis are survival, employment, and innovation as reflected by patenting activity and patent citations.[5]

2.2 Which Firms do Venture Capitalists Select?

Many venture capitalists typically identify and fund startups very early during their formation. Very few of startups grow quickly and become successful firms, while most grow more slowly or fail. On average, however, venture capitalist expertise results in the selection of startups that are more promising than the typical startup in the economy.

The age distribution of startups at the time of first funding is in Figure 2. Age is measured by the number of years since the first observation (birth) of a startup as an employer business in the LBD—negative age values correspond to cases where a startup receives funding even before it is first observed as an employer business. The striking feature of Figure 2 is the spike at age zero. About 42 percent of VC-funded firms receive their first funding in their first year as an employer business. There is also a large fraction of startups that receive their first funding before they hire their first employee (age < 0).[6] [7]

Figure 3 plots the probability of receiving VC funding versus startup employment growth quintiles during the first three years of a startup.[8] The probability of funding jumps nearly

[4]For details on identifying IPOs and M&As, see Section 10. We thank Joonkyu Choi and Javier Miranda for their assistance on identifying M&A in the Census microdata based on their methodology.

[5]The results using revenue are briefly mentioned as an alternative to employment when appropriate. For firm size, employment is the preferred measure in the analysis, since annual revenue data is available starting only in 1997.

[6]These firms most likely receive VC funding at a stage when they are still non-employers; i.e., businesses with no payroll/employees.

[7]For the rest of the analysis, startups that receive VC funding before age zero are treated as having received funding at age zero. This approach is necessary because several key variables (e.g. employment and patenting) for the subsequent analysis are only available only for employer firms starting at age zero.

[8]The employment growth rate for the first three years of a firm's life-cycle is measured as $g = 2 \times (emp_{t+2} - emp_t)/(emp_{t+2} + emp_t)$, where t is the year a startup is first observed as an employer business. The growth measure takes a value in the range $[-2, 2]$, where the endpoints 2 and -2 correspond, respectively, to the growth rates of startups that start with zero employment, and those startups that fail by $t + 2$ (and

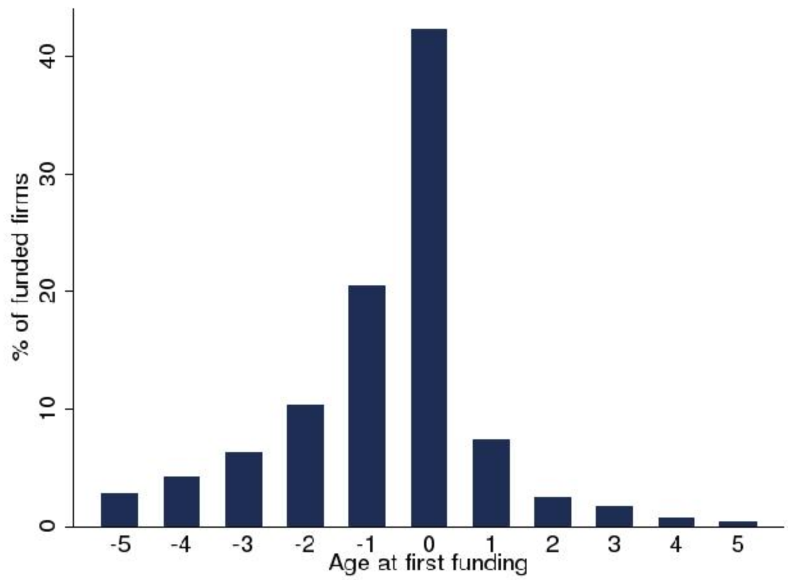

Figure 2: The distribution of firm age in the year of first VC funding. Age 0 corresponds to the first time that a startup is observed as a business with employee(s). Ages -5 to -1 correspond to startups for which the year of first VC funding occurs before the first observation of the startup as an employer.

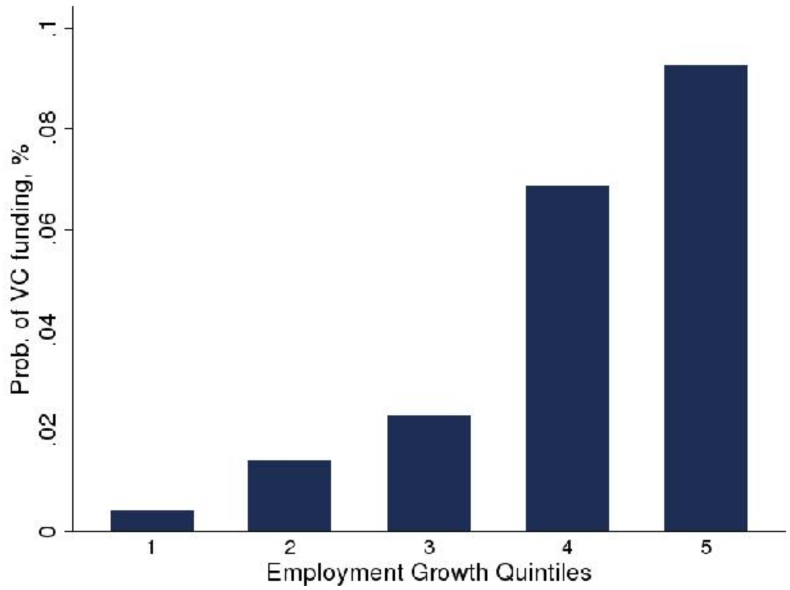

Figure 3: Firm employment growth and the probability of VC funding. Employment growth is the 3-year growth rate for a startup, starting from the first year of observation as an employer.

two hundred-fold as one moves from the lowest quintile to the highest. This sharp rise suggests that venture capitalists tend to select firms that exhibit relatively high growth in their early phases. (A similar conclusion emerges when growth is measured in terms of revenue instead of employment.)

Figure 4 suggests that VC funding is concentrated in startups with high innovative potential. The figure plots the relationship between the startup's probability of receiving VC funding within five years of the first patent application and the quality of that patent, measured by citations. The probability of receiving VC funding for firms in the bottom quintile of the patent quality distribution is only about 1 percent, but rises to nearly 6 percent in the top quintile.

Table 1 presents some statistics on startups in the year of VC funding. The average employment at funding is normalized to one for VC-funded startups that are in the top decile of firm employment distribution in the year the first funding is received. For non-

hence have zero employment then).

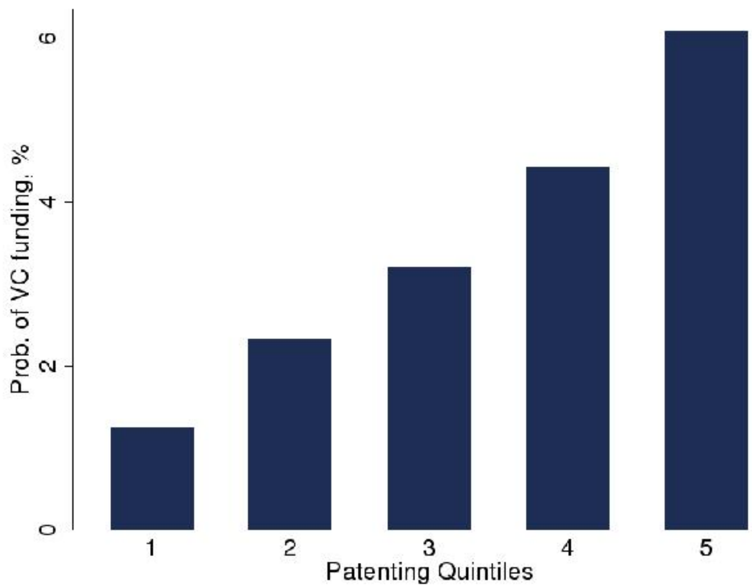

Figure 4: The quality of a firm's first patent and the probability of VC funding. The figure is conditional on having a patent. Patenting quintiles give the quality distribution for a startup's first patent. Patent quality is measured by the number of citations demeaned by the citations in its patent class, excluding self-citations. The probability of VC funding is measured within 5 years of the first patent.

funded startups, the hypothetical age at first funding is set to the average age at which a VC-funded firm receives funding.[9] Table 1 indicates that at the time of funding the average employment of VC-funded firms is about 3 times larger than that of non-funded firms.[10] VC-funded startups in the top decile ($\geq$ 90th percentile) of the employment distribution are about 43 percent larger than the non-funded firms that are also in the top decile. There is not a big difference, however, in the average employment of funded versus non-funded firms in the bottom 90 percent ($\leq$ 90th).[11]

Overall, the patterns in Figures 3, 4, and Table 1 suggest the presence of a high degree of selection in VC funding. Venture capitalists choose startups that are on average larger, and exhibit higher early-stage employment growth and innovation. Furthermore, venture capitalists select such startups very early in their life-cycles, as indicated by the firm age distribution at the time of first VC funding. This timing is consistent with venture capitalists' reliance on their experience and expertise to assess the promise of a startup almost immediately upon its inception.[12]

---

[9] The average age at VC funding is calculated using the age distribution of VC-funded firms in the year of first VC funding, assuming a VC-funding age of zero for firms that receive VC funding before age zero.

[10] The ratio of the average sizes is given by $r = (E_{VC}/N_{VC})/(E_{Non\text{-}VC}/N_{Non\text{-}VC})$, which can be written as

$$\frac{(E_{VC}/E)/(N_{VC}/N)}{(E_{Non\text{-}VC}/E)/(N_{Non\text{-}VC}/N)},$$

where E is the total employment and N is the total number of firms. From Table 1, $E_{VC}/E = 0.25\% + 0.04\% = 0.3\%$, $N_{VC}/N = 0.03\% + 0.07\% = 0.1\%$, $E_{Non\text{-}VC}/E = 57.8\% + 41.9\% = 99.7\%$, and $N_{Non\text{-}VC}/N = 10.7\% + 89.2\% = 99.9\%$. Therefore, the ratio r is given by $(0.3/0.1)/(99.7/99.9) = 3.0$.

[11] A caveat to these comparisons of startup size at the time of VC funding is that most startups at this stage are very small, and there is not a very large variation in size across firms within top 10 percent or within bottom 90 percent. Firm sizes diverge substantially by type as startups grow and selection takes place, as discussed below.

[12] Anecdotal evidence suggests that venture capitalists may even target early stage projects with no business plan at all. See, for example, the interview "Venture capitalism: Investing today in the companies of tomorrow" by Ryssdall and Hollenhorst in NPR's Marketplace, which features venture capitalist Ann Miuro-Ko's investment of $1M in an early project that consists of merely a few math papers written by a Stanford professor. (https://www.marketplace.org/2018/08/28/tech/corner-office/venture-capitalism-investing-today-ideas-tomorrow)

Statistics for Firms—Year of First VC Funding

	Share of (%)						Average	
	Firms		*Empl*		*Firms by funding type*		*Empl (Rel)*	
Type	$\geq$ 90th	$\leq$ 90th	$\geq$ 90th	$\leq$ 90th	$\geq$ 90th	$\leq$ 90th	$\geq$ 90th	$\leq$ 90th
VC funded	0.03	0.07	0.25	0.04	32.6	67.4	1.00	0.05
Non-VC funded	10.7	89.2	57.8	41.9	10.7	89.4	0.70	0.06

Table 1: A firm's type refers to whether or not it is in the top decile (90th percentile) for employment at the time of first VC funding. Average employment is measured with respect to the VC-funded firms in the top decile. Non-VC-funded firms are assigned a hypothetical age of funding equal to the average age at first funding for VC-funded firms. All figures are rounded for disclosure purposes.

2.3 How Much do Venture Capitalists Matter?

Next, consider the effects of venture capitalist involvement on the post-funding evolution of a startup.[13] To assess the magnitude of these effects, the selection of startups by venture capitalists must be taken into account. To control for selection based on observables, a group of non-VC-funded firms that closely match VC-funded firms on a number of key characteristics are identified. Because there is a large number of potential matches in the constructed dataset, it is possible to find exact matches for any VC-funded firm based on many of the key characteristics. Each firm that received first VC funding in a given year is matched with a non-VC-funded firm in the same year that is of the same age, in the same narrowly defined employment bin, operates in the same 4-digit NAICS industry, has the same multi-unit/single-unit firm designation, and is located in the same state.[14] These match criteria are quite stringent, and ensure that many key observable characteristics are

[13] One caveat is that no direct measures of VC involvement are available, such as the presence of a venture capitalist on the management board of a funded firm or the amount of time spent by a venture capitalist on mentoring an entrepreneur.

[14] The main firm size measure, employment, cannot always be matched exactly conditional on other match criteria, especially for relatively large VC-funded firms. Instead, a coarser matching based on employment bins is considered.

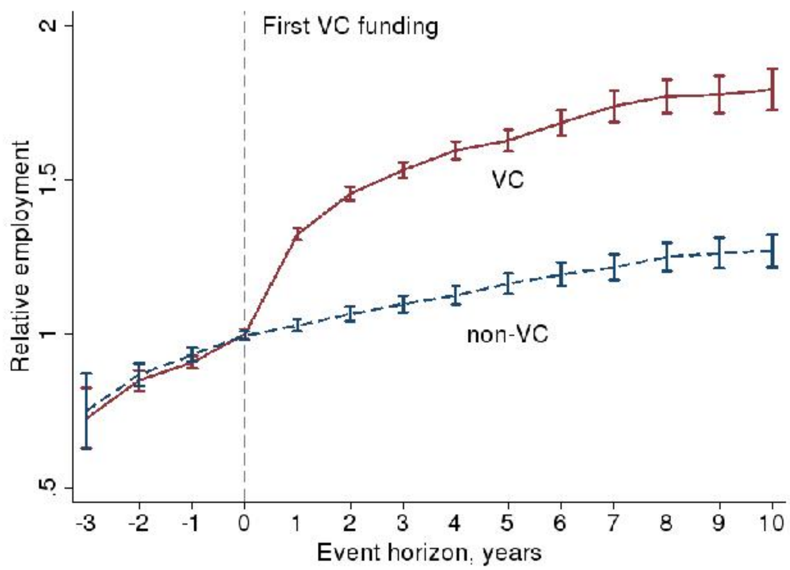

Figure 5: The evolution of average employment before and after first VC funding: VC-funded firms versus non-VC-funded controls. The observations are relative to average employment (normalized to 1) for VC-funded startups in the year of first VC funding, $t = 0$.

closely accounted for. The treated and control groups are used to perform an event study that considers employment and the quality-adjusted patent stock as the key outcome variables—see Appendix 10.3.4 for further details on the construction of event study samples.

Figure 5 plots the evolution of (ln) average employment for VC-funded firms and their matched counterparts. The data points in Figure 5 are relative to the (ln) average employment of VC-funded firms at the time of first VC funding ($t = 0$), which is normalized to 1. The two types of firms are followed up to 10 years after they receive VC funding, and up to 3 years before that point—that is, only the firms that receive VC at ages 1, 2 or 3 are included.[15] The horizon of 10 years is chosen to approximate the duration of VC contracts, which is about a decade.

[15]Firms are followed until the point where either they fail or get involved in an IPO, a merger, or an acquisition. That is, the samples are not balanced panels; the evolution of average employment is influenced by selection due to firm failure, IPO and M&A activity. In addition, firms that receive VC funding at age zero are excluded, since these firms have no observations for the pre-VC period for a proper event study. Last, all firms that report zero employment are excluded from the analysis for both groups.

The two groups exhibit a very similar trajectory of average employment before VC funding, indicating that the matching algorithm does a decent job of accounting for the pre-VC-funding trends. The funded firms, however, grow, on average, much more following the funding. Average employment increases by approximately 475 percent by the end of the horizon for VC-funded firms, whereas growth is much more modest for the control group (about 230 percent). Based on these differential growth rates, the treatment by venture capitalists accounts for up to more than half of the growth experienced by VC-funded startups.[16]

The failure rate of VC-funded firms in the 10 years following funding is slightly larger than that of the non-funded firms (36 percent versus 34 percent). However, the growth rates of VC-funded firms, conditional on survival, exhibit considerable divergence. These rates are also very different from those of the non-VC-funded startups. The kernel density estimates for the growth rates of employment conditional on survival 10 years after VC funding are plotted in Figure 6 for VC-funded and non-VC-funded firms.[17] The growth-rate distribution for VC-funded firms is negatively skewed, with a large mass of firms beyond the growth rate of 100 percent, and a peak density around the growth rate of 200 percent. In contrast, non-VC-funded firms have a much more symmetric growth-rate distribution that is centered slightly above zero.[18] The dispersion of growth rates is also much higher for VC-funded firms. VC involvement is associated with more dispersed, but on average much higher, employment growth rates conditional on survival.[19]

The evolution of innovative activity, as measured by the (ln) quality-adjusted patent stock, is shown in Figure 7.[20] As in Figure 5, the observations are relative to the (ln)

[16] Although not reported, the patterns are largely similar if revenue is considered as a measure of firm size. Note that revenue was not used as a variable in selecting the control group.

[17] This is done in a manner similar to footnote 8.

[18] The relatively low median 10-year growth rate for non-VC-funded firms is consistent with earlier findings which indicate a very low average growth rate for startups in the U.S. economy—see, for example, Decker et al. (2016) and Choi (2018).

[19] The growth patterns are similar if revenue is used instead of employment as the outcome measure.

[20] This figure only includes firms that receive VC funding at ages 1 and 2 only, compared to Figure 5, which also includes firms that receive funding at age 3. The reason is that the control group for firms that receive VC funding at age 3 is much smaller and only contain a small number of firms that have patents, leading to a very large confidence interval for the average patent stock.

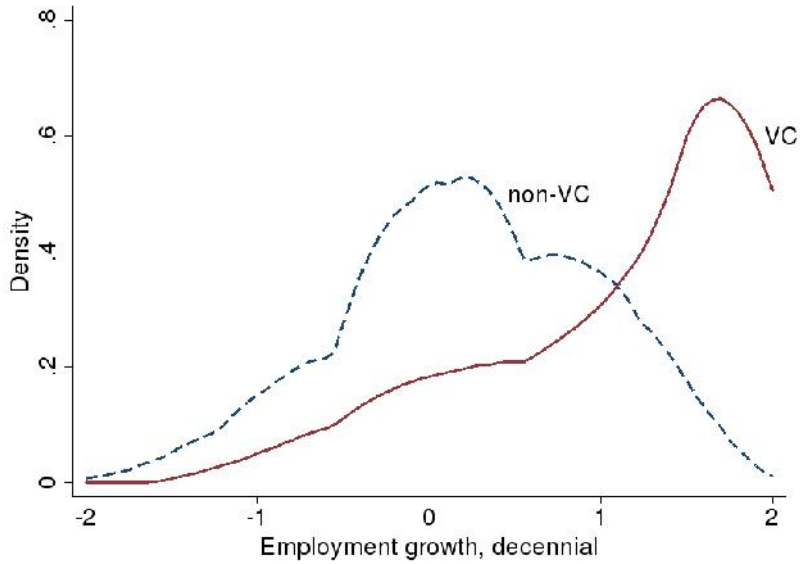

Figure 6: The distribution of the employment growth rate, measured 10 years after first VC-funding, for both VC-funded and non-VC-funded firms. The figure is conditional on survival.

average patent stock of VC-funded firms at the time of funding ($t = 0$), which is normalized to 1. Pre-VC trends for the two groups are similar, again indicating a close match based on the same criteria used in the employment case—even though patent stock similarity is not one of the criteria. The post-VC trends indicate that VC-funded firms experience, on average, a much higher growth in their patent stock. The VC-funded firms' average patent stocks grows by about $1,100$ percent, as opposed to about 440 percent for the control group. These rates suggest that up to 60 percent of the growth in the patent stock can be attributed to treatment by venture capitalists.

The results indicate strong post-funding effects associated with VC involvement. VC-backed firms achieve on average better outcomes in terms of employment, innovation, and revenue, after controlling for selection based on observables. Moreover, VC-backed startups are much more likely to make it to the right tail of firm employment and innovation distributions, compared with non-VC-funded ones. Some additional statistics on the superior performance of VC-backed, relative to non-VC-backed, firms are presented in Section 6. Overall, venture capitalist involvement appears to be an important mechanism that generates

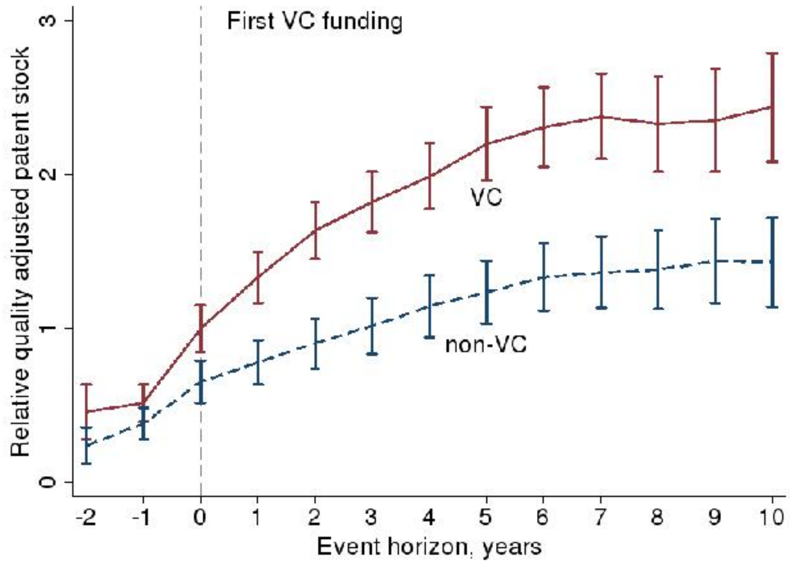

Figure 7: The evolution of the average patent stock before and after first VC funding: VC-funded firms versus non-VC-funded controls The observations are relative to the average patent stock (normalized to 1) for VC-funded firms in the year of VC funding, $t = 0$.

successful firms and high degree of skewness in firm outcome distributions.

2.4 Assortative Matching between Venture Capitalists and Startups

Are the effects of VC involvement documented in the previous sections larger for more able and more experienced venture capitalists? To answer this question, venture capitalists in the sample are divided into two groups: low- versus high-quality. A venture capitalist is labelled "high-quality" if it is in the 90th percentile of the distribution over the total number of deals (total number of funding rounds) by a venture capitalist over the entire sample period, and "low-quality" otherwise.[21] Alternative definitions of quality are also considered, as discussed

[21]It is common for VC-funded firms to receive funding from multiple venture capitalists. Based on the venture capitalist quality/experience measure constructed, if 50 percent or more of a firm's funders are of high quality, the firm is considered to be funded by high-quality venture capitalists. If instead, more than 50 percent of its funders are of low quality, it is considered to be a firm funded by low-quality venture capitalists.

below. VC-funded firms are divided into two mutually exclusive groups: those funded by high-quality venture capitalists, and those funded by low-quality ones.[22]

Figure 8 plots the evolution of (ln) average employment of firms by venture capitalist quality. Observations are relative to the (ln) average employment level of startups funded by high-quality venture capitalists measured at the time of first funding ($t = 0$), which is normalized to 1. In the pre-funding periods ($t < 0$), the trajectories for the two groups are similar, but at $t = 0$ average employment is higher for firms funded by high-quality venture capitalists. In the years following VC funding ($t > 0$), firms funded by high-quality venture capitalists grow significantly more on average. Following a notable jump at $t = 1$, the gap between the trajectories continues to open up. By the end of the horizon, average employment grows by about 400 percent in the high-quality group, and by about 320 percent in the low-quality group.

A similar picture emerges when the event study is repeated for the (ln) average quality-adjusted patent stock, as shown in Figure 9.[23] While the difference between the firms funded by high-quality versus low-quality venture capitalists is small for the periods before funding ($t < 0$), there is a gap at $t = 0$, which grows over time after funding. By the end of the horizon, the average patent stock grows by nearly 50-fold for the high-quality group, and only about 19-fold for the low-quality group.[24]

There is also evidence that the quality of a venture capitalist matters beyond just the amount of funding provided. Unobserved aspects of venture capitalist involvement in a startup, such as mentoring or monitoring, appear to matter for the post-funding performance

[22] The set of VC-funded firms considered in the analysis to follow is larger than the one used to generate Figures 5 and 7, because the latter consider only the VC-funded firms for which close matches within the set of non-funded firms are available. In addition, some VC-funded firms have missing funding information, which generates additional differences from the samples used in Figures 5 and 7.

[23] Note that the pre-VC period extends to 3 periods in this figure, compared to only 2 periods in the analysis associated with Figure 7. The reason is that in the latter analysis the control group is very small and does not include a large enough number of patenting firms for $t < -2$.

[24] The findings in Figures 8 and 9 are qualitatively similar when revenue is considered as an alternative outcome. Many of the results are also similar when alternative measures of quality for venture capitalists are considered—for instance, when quality is measured by total all-time funding by a venture capitalist, by average funding per year, or average number of deals per year.

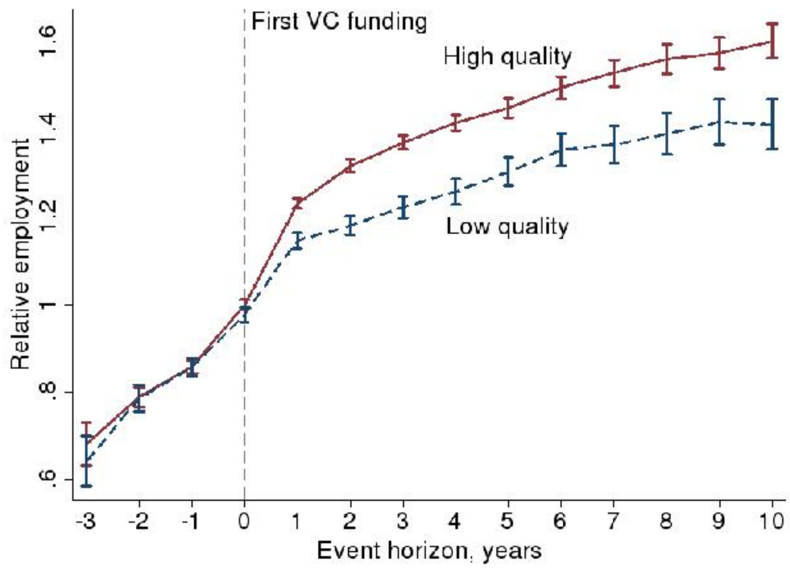

Figure 8: The evolution of average employment before and after first VC funding, by the quality of the venture capitalist. The observations are relative to the average employment (normalized to 1) for startups funded by high-quality venture capitalists in the year of first VC funding, $t = 0$.

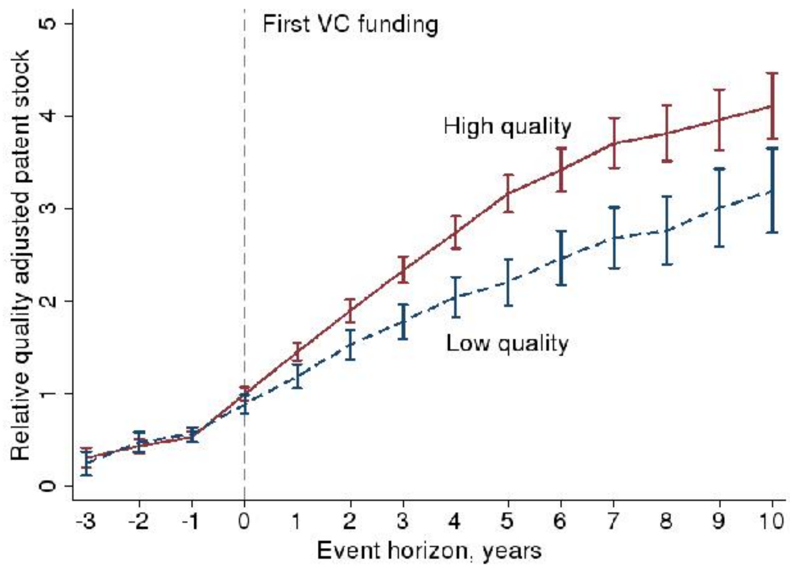

Figure 9: The evolution of the average patent stock before and after first VC funding, by the quality of the venture capitalist. The observations are relative to the average patent stock (normalized to 1) for startups funded by high-quality venture capitalists in the year of first VC funding, $t = 0$.

of funded startups. To quantify the effects of such unobserved aspects, the employment of a startup 10 years after funding is projected on the initial amount of VC funding and a dummy variable that indicates backing by a high-quality venture capitalist—a variable that controls for all venture capitalist-specific fixed factors in a reduced form. Employment and age at the time of funding are also included to control for selection effects. Because some startups fail or end up in M&A before the end of the 10-year horizon, years from first funding is also added to control for the length of venture capitalist involvement. Finally, industry and year fixed effects are included to control for variation in VC selection of sectors and economy-wide factors that can influence firm outcomes.

Results from alternative specifications are shown in Table 2. In all specifications, the funding level has positive and highly significant coefficients, which imply about a 0.35 to 0.41 percent increase in ultimate employment corresponding to a 1 percent increase in funding. Startups that are larger at the time of funding also end up being even larger at the end of the treatment, while startups that receive funding at a later age do worse. The key coefficient estimates, associated with the dummy for high-quality venture capitalists, all indicate a positive association between quality and ultimate employment. In particular, controlling for the funding level and other observables (specification 4), startups backed by a high-quality venture capitalist are estimated to end up with approximately 32 percent larger employment.[25] This finding suggests better outcomes for firms backed by venture capitalists of higher quality, beyond simply the funding they provide.[26]

Overall, empirical findings suggest both strong selection and treatment effects associated with venture capitalist involvement. VC funding is concentrated in firms that are younger, larger and have better growth and innovation prospects. Post-funding outcomes of VC-funded firms in terms of firm size and innovation far exceeds those of non-VC-funded firms that are otherwise similar to VC-funded firms in key observable characteristics. Moreover, it

[25] The estimated effect is given by $e^{0.274} - 1 \simeq 0.32$.

[26] The conclusions from Table 2 also apply to alternative measures of quality for a venture capitalist. For various quality measures used, there is a positive association between the ultimate employment of a startup and the quality of the venture capitalist associated with that startup.

Employment Outcomes and Quality of Venture Capitalists

	Employment at 10 years from funding (ln)			
Independent Variables	1	2	3	4
High-quality funder (dummy)	0.253***	0.260***	0.292***	0.274***
	(0.030)	(0.029)	(0.028)	(0.028)
Equity at first funding (ln)	0.414***	0.368***	0.352***	0.347***
	(0.013)	(0.012)	(0.012)	(0.012)
Employment at first funding (ln)		0.267***	0.278***	0.249***
		(0.008)	(0.008)	(0.008)
Age at first funding			−0.074***	−0.082***
			(0.006)	(0.006)
Years since first funding				0.146***
				(0.004)
Industry F.E.	Yes	Yes	Yes	Yes
Year F.E.	Yes	Yes	Yes	Yes
N	14,000	14,000	14,000	14,000
R^2	0.17	0.23	0.24	0.30

Table 2: Notes: Standard errors clustered by industry are in parentheses. N is rounded to avoid disclosure of confidential information. N is rounded to avoid the disclosure of confidential information. The high-quality funder dummy equals 1, if the venture capitalist is in the top decile of the distribution of the total number of deals made by a venture capitalist. The omitted category is the low-quality funder. The regressions include a constant, and industry and year fixed effects. All figures are rounded for disclosure purposes.

is not just the funding level that matters—startups that are backed by high-quality venture capitalists achieve better outcomes even after controlling for funding. The next section proposes a model that can account for many of the stylized facts presented so far.

3 The Model

Every period s new entrepreneurs approach financiers to acquire funding for startups, where s is exogenous. Entrepreneurs come in two *ex ante* types, $i \in \{h, l\}$, namely high or low. High-type entrepreneurs are more likely to develop a successful project, have a higher level of synergy with a financier, and attain a higher level of productivity, other things equal, than low-type entrepreneurs. There are 2 types of financiers, $f \in \{v, n\}$; namely, venture capitalists, v, and non-venture capitalists, n, corresponding to traditional financial institutions, such as banks. Out of the s new startups, s^n are financed by banks and s^v by venture capitalists. A venture capitalist nurtures startups by providing both expertise and funding. A traditional financial institution offers a much more limited amount of expertise. In the equilibrium being modeled, $s^v < s$ so that not all new entrepreneurs can be financed by venture capitalists; some must go to traditional financial institutions. The analysis will focus on balanced growth paths. Specifically, aggregate productivity, $\mathbf{x}$, will grow at some endogenous gross rate, $g_{\mathbf{x}} > 1$, which in turn results in wages, w, growing at another constant endogenous rate, $g_w > 1$.

Funded startups have two sources of uncertainty. First, the startup may fail. Second, even if it succeeds, the startup may enter either a high or low *ex post* state for production, represented by $j \in \{h, l\}$. There are synergies between an entrepreneur and a financier. The synergy between a financier of type f and a type-(i, j) entrepreneur, who starts off in ex ante state i and who then transits to ex post state j, is represented by z_{ij}^f, for $f = v, n$ and $i, j = h, l$. Since venture capitalists provide more nurturing for startups, it transpires that $z_{ij}^v > z_{ij}^n$ for $i, j = h, l$. Additionally, because the high states are better than the low ones, for a given type of financing, $z_{hj}^f > z_{lj}^f$ and $z_{ih}^f > z_{il}^f$, for $f = v, n$.

The matching process between entrepreneurs and financiers is random. The fraction μ_h^v of the projects financed by the venture capitalist will be with ex ante high-type entrepreneurs, Ex ante Matching

Financing	Entrepreneurs, fraction	
	High	Low
VC Funded	μ_h^v	$\mu_l^v = 1 - \mu_h^v$
Non-VC Funded	μ_h^n	$\mu_l^n = 1 - \mu_h^n$

Table 3: By construction the rows sum to one. The number of VC-funded startups in the ex ante high are low states are $s^v \mu_h^v$ and $s^v \mu_l^v$, while the corresponding numbers for non-VC-funded startups read $s^n \mu_h^n$ and $s^n \mu_l^n$.

while the remaining fraction, $\mu_l^v = 1 - \mu_h^v$, will be with low-type ones. The corresponding fractions for non-VC funding will be given by μ_h^n and $\mu_l^n = 1 - \mu_h^n$. So, matching is summarized by Table 3. There is positive assortative matching when $\mu_h^v > \mu_h^n$, which turns out to be the case in the U.S. data.

3.1 Floated Ventures

The output, o_{ij}^f, of a successful ex post type-(f, i, j) startup is

$$o_{ij}^f = (x_{ij}^f z_{ij}^f)^\zeta (k_{ij}^f)^\kappa (l_{ij}^f)^\lambda, \text{ with } \zeta + \kappa + \lambda = 1, \text{ for } f = v, n \text{ and } i, j = h, l. \tag{1}$$

where k_{ij}^f and l_{ij}^f are the amounts of capital and labor used in production, x_{ij}^f is the firm's level of productivity, and z_{ij}^f is the synergy level between the entrepreneur and financier. A producing firm can freely rent capital at the rate r and hire labor at the wage rate w. It does so to maximize profits. The rental price for capital incorporates a depreciation rate of $1 - \mathfrak{d}$. The profit function associated with this maximization problem is

$$x_{ij}^f z_{ij}^f (1 - \kappa - \lambda)[(\frac{\kappa}{r})^\kappa (\frac{\lambda}{w})^\lambda]^{1/\zeta} = \max_{k_{ij}^f, l_{ij}^f} \{o_{ij}^f - r k_{ij}^f - w l_{ij}^f : \text{ s.t. } (1)\}, \tag{2}$$

for $f = v, n$ and $i, j = h, l$. Along a balanced growth path wages will grow at some (endogenous gross) rate, g_w. As can be seen from the profit function, rising wages operate to reduce the profits of an enterprise as it ages. The present value of the profits from an IPO (or an acquisition) in the *current* period is

$$
\begin{aligned}
& x_{ij}^f z_{ij}^f \sum_{t=0}^{\infty} (\frac{\delta \mathbf{s}}{g_w^{\lambda/\zeta}})^t (1-\kappa-\lambda)[(\frac{\kappa}{r})^\kappa (\frac{\lambda}{w})^\lambda]^{1/\zeta} \\
\equiv \; & x_{ij}^f z_{ij}^f I(\cdot), \text{ for } f = v, n \text{ and } i, j = h, l,
\end{aligned} \tag{3}
$$

where $\mathbf{s}$ is the survival rate for a firm. Note that the value of an IPO is linear in $x_{ij}^f z_{ij}^f$.

3.2 Startups

The productivity level of a type-(f, i, j) startup, x_{ij}^f, is endogenous. The more ambitious a new enterprise is at the time of inception, or the higher productivity, x_{ij}^f, is relative to the average productivity of currently producing firms, denoted by $\boldsymbol{x}$, the greater is the initial research cost. The state-specific research cost for a startup is specified by the function

$$
R(\frac{x_{ij}^f}{\mathbf{x}}) = w(\frac{x_{ij}^f}{\mathbf{x}})^\iota / (\iota \chi_{ij}^f), \text{ with } \iota > 1, \text{ and for } f = v, n \text{ and } i, j = h, l.
$$

The likelihood that a startup will be successful depends upon the amount of development funding it receives. Let σ_{ij}^f denote the odds that a type-(f, i, j) venture will be successful. These odds can be secured with a funding level of $D(\sigma_{ij}^f)$, where

$$
D(\sigma_{ij}^f) = w(\frac{1}{1-\sigma_{ij}^f} - 1)\sigma_{ij}^f / \psi_{ij}^f, \text{ for } f = v, n \text{ and } i, j = h, l.
$$

Note that R&D is done separately for each ex post state, $j = h, l$, before the realization of the state is known. The entrepreneur must seek outside funding to cover these R&D startup costs, $R(x_{ih}^f/\mathbf{x}) + D(\sigma_{ih}^f) + R(x_{il}^f/\mathbf{x}) + D(\sigma_{il}^f)$, where again there are two mutually exclusive sources for startup funding, $f = v, n$.

4 Decision Problems

4.1 VC-Financed Startups

Suppose that an ex ante type-i entrepreneur receives funding and mentoring from a venture capitalist. The entrepreneur and the venture capitalist sign a contract, which stipulates that the entrepreneur must pay the venture capitalist p^v_{ih} if the project is successful in the high state and p^v_{il} if it lands in the low state. The probabilities of success in the high and low states, σ^f_{ih} and σ^f_{il}, are also specified by the contract, as are the productivities, x^v_{ih} and x^v_{il}. The terms of the agreement are prescribed by the following Nash Bargaining problem

$$\begin{aligned}\max_{\sigma^v_{ih},\sigma^v_{il},x^v_{ih},x^v_{il},p^v_{ih},p^v_{il}} &[\sigma^v_{ih}\delta x^v_{ih}z^v_{ih}I(\cdot') + \sigma^v_{il}\delta x^v_{il}z^v_{il}I(\cdot') - \sigma^v_{ih}\delta p^v_{ih} - \sigma^v_{il}\delta p^v_{il} - E_i] \\ &\times [\sigma^v_{ih}\delta p^v_{ih} + \sigma^v_{il}\delta p^v_{il} - R(\frac{x^v_{ih}}{\mathbf{x}}) - R(\frac{x^v_{il}}{\mathbf{x}}) - D(\sigma^v_{ih}) - D(\sigma^v_{il}) - V],\end{aligned} \tag{P1}$$

where E_i and V are the threat points of the entrepreneur and the venture capitalist, respectively. Here δ is the market discount factor, or the reciprocal of the risk-free gross interest rate on saving and lending by consumers. This is used to discount the payoffs to both parties for each ex post state $j = h, l$. The term in brackets on the first line presents the expected reward to the entrepreneur, net of his threat point, while the term on the second line gives the same thing for the venture capitalist.

The upshot of the above Nash Bargaining problem is the following three first-order conditions:

$$\delta x^v_{ij}z^v_{ij}I(\cdot') = D_1(\sigma^v_{ij}), \text{ for } j = h, l, \tag{4}$$

$$\delta \sigma^v_{ij}z^v_{ij}I(\cdot') = R_1(\frac{x^v_{ij}}{\mathbf{x}})/\mathbf{x}, \text{ for } j = h, l, \tag{5}$$

and

$$\begin{aligned}\sigma^v_{ih}\delta p^v_{ih} + \sigma^v_{il}\delta p^v_{il} = R(\frac{x^v_{ih}}{\mathbf{x}}) + R(\frac{x^v_{il}}{\mathbf{x}}) + D(\sigma^v_{ih}) + D(\sigma^v_{il}) \\ + \frac{1}{2}[\sigma^v_{ih}\delta x^v_{ih}z^v_{ih}I(\cdot') + \sigma^v_{il}\delta x^v_{il}z^v_{il}I(\cdot') \\ - R(\frac{x^v_{ih}}{\mathbf{x}}) - R(\frac{x^v_{il}}{\mathbf{x}}) - D(\sigma^v_{ih}) - D(\sigma^v_{il}) + V - E_i].\end{aligned}$$

The first condition implies that the probability of success for each state j is chosen in the first-best manner. A marginal increase in the state-j success probability, σ^v_{ij}, raises the discounted value of an IPO by $\delta x^v_{ij}z^v_{ij}I(\cdot')$, which is the marginal benefit. Development costs move up by $D_1(\sigma^v_{ij})$, which represents the marginal cost. So, the marginal benefit from increasing σ^v_{ij} equals its marginal cost. The next condition sets for each state j the discounted expected marginal benefit from an increment in productivity, $\delta\sigma_{ij}^v z_{ij}^v I(\cdot')$, equal to the marginal cost of research, $R_1(x_{ij}^v/\mathbf{x})/\mathbf{x}$. Thus, x_{ij}^v is also set in a first best manner. The third condition specifies the venture capitalist's expected discounted payment, $\sigma_{ih}^v\delta p_{ih}^v + \sigma_{il}^v\delta p_{il}^v$. It is equal to his R&D startup costs, $R(x_{ih}^v/\mathbf{x}) + R(x_{il}^v/\mathbf{x}) + D(\sigma_{ih}^v) + D(\sigma_{il}^v)$, plus one half of the expected surplus, net of the difference in threat points, or $\sigma_{ih}^v\delta x_{ih}^v z_{ih}^v I(\cdot') + \sigma_{il}^v\delta x_{il}^v z_{il}^v I(\cdot') - R(x_{ih}^v/\mathbf{x}) - R(x_{il}^v/\mathbf{x}) - D(\sigma_{ih}^v) - D(\sigma_{il}^v) + V - E_i$. The upshot of the first two conditions is summarized by the following lemma.

Lemma 1 *(The Impact of Synergy) The higher the ex post state j synergy, z_{ij}^v, between an ex ante type-i entrepreneur and the venture capitalist, the more ambitious a project will be as reflected by:*

1. *a higher level of productivity, x_{ij}^v;*
2. *a higher likelihood of success, σ_{ij}^v; and*
3. *a higher level of funding, $R(x_{ij}^v/\mathbf{x}) + D(\sigma_{ij}^v)$,*

for $j = h, l$.

Proof. See Appendix 11. ■

Now, an ex ante type-i entrepreneur can always go to a bank for funding. So, his threat point, E_i, is given by maximized value of the bank startup problem (P2) presented below. Now, suppose that a venture capitalist has only one draw in the matching scenario. Then he will have no outside option. Thus, his threat point is given by $V = 0$. Therefore, the venture capitalist's expected profits from a contract with ex ante type-i entrepreneur, V_i, are

$$V_i = [\sigma_{ih}^v\delta x_{ih}^v z_{ih}^v I(\cdot') + \sigma_{il}^v\delta x_{il}^v z_{il}^v I(\cdot') - R(x_{ih}^v/\mathbf{x}) - R(x_{il}^v/\mathbf{x}) - D(\sigma_{ih}^v) - D(\sigma_{il}^v) - E_i]/2, \quad (6)$$

which represents half of the surplus from the enterprise net of the entrepreneur's threat point.

4.2 Non-VC-Financed Startups

A banker lends to an ex ante type-i entrepreneur at the market discount factor i_i, which takes into account that the startup might fail with probability $1 - \sigma_{ih}^n - \sigma_{il}^n$. When a startup fails, the entrepreneur defaults on the loan's interest and principal, $[D(\sigma_{ih}^n) + R(x_{ih}^n/\mathbf{x}) + D(\sigma_{il}^n) + R(x_{il}^n/\mathbf{x})]/i_i$. The gross interest rate charged by banks, $1/i_i$, will be determined by a zero-profit constraint for bankers. If bankers can raise funds at the risk-free rate $1/\delta$, then

$$\frac{\sigma_{ih}^n + \sigma_{il}^n}{i_i} = \frac{1}{\delta} \text{ or } \frac{1}{i_i} = \frac{1}{\delta(\sigma_{ih}^n + \sigma_{il}^n)}. \tag{7}$$

Therefore, the gross interest rate charged by banks, $1/i_i$, is decreasing in the success probability, $\sigma_{ih}^n + \sigma_{il}^n$. Equivalently, the gross interest rate is increasing in the failure rate, $1 - \sigma_{ih}^n - \sigma_{il}^n$. Clearly, $1/i_i > 1/\delta$ or $i_i < \delta$.

Now, the entrepreneur's maximization problem is

$$\begin{aligned} E_i \;=\; & \max_{\sigma_{ih}^n, \sigma_{il}^n, x_{ih}^n, x_{il}^n} \delta\{\sigma_{ih}^n x_{ih}^n z_{ih}^n I(\cdot') + \sigma_{il}^n x_{il}^n z_{il}^n I(\cdot') \\ & -(\sigma_{ih}^n + \sigma_{il}^n)[R(\frac{x_{ih}^n}{\mathbf{x}}) + R(\frac{x_{il}^n}{\mathbf{x}}) + D(\sigma_{ih}^n) + D(\sigma_{il}^n)]/i_i\}, \end{aligned} \tag{P2}$$

subject to (7). Here the entrepreneur takes into account how his plan for the startup influences the gross interest rate that he pays. The two sets of first-order conditions associated with the above maximization problem are:

$$\delta x_{ij}^n z_{ij}^n I(\cdot') = D_1(\sigma_{ij}^n), \text{ for } j = h, l, \tag{8}$$

and

$$\delta \sigma_{ij}^n z_{ij}^n I(\cdot') = R_1(\frac{x_{ij}^n}{\mathbf{x}})/\mathbf{x}, \text{ for } j = h, l. \tag{9}$$

These two efficiency conditions for σ_{ij}^n and x_{ij}^n have the same form as (4) and (5). The values for σ_{ij}^n and x_{ij}^n will differ from σ_{ij}^v and x_{ij}^v since $z_{ij}^n < z_{ij}^v$. By appealing to Lemma 1, it is clear that, because $z_{ij}^n < z_{ij}^v$, non-VC-financed startups will have lower levels of productivity, x_{ij}^n, success, σ_{ij}^n, and funding, $R(x_{ij}^n/\mathbf{x}) + D(\sigma_{ij}^n)$.

5 Balanced Growth

To start with, denote the gross growth rate for aggregate productivity, or $\mathbf{x}$, by $g_{\mathbf{x}} \equiv \mathbf{x}'/\mathbf{x}$. Define the average level of productivity for a type-f financed venture in the ex post high state by

$$x_h^f = (\mu_h^f \sigma_{hh}^f x_{hh}^f + \mu_l^f \sigma_{lh}^f x_{lh}^f)/(\mu_h^f \sigma_{hh}^f + \mu_l^f \sigma_{lh}^f), \text{ for } f = n, v.$$

The fractions of all successful startups that are of ex post type (f, h) are

$$\nu_h^f = \frac{s^f(\mu_h^f \sigma_{hh}^f + \mu_l^f \sigma_{lh}^f)}{s^v(\mu_h^v \sigma_{hh}^v + \mu_l^v \sigma_{lh}^v) + s^n(\mu_h^n \sigma_{hh}^n + \mu_l^n \sigma_{lh}^n) + s^v(\mu_h^v \sigma_{hl}^v + \mu_l^v \sigma_{ll}^v) + s^n(\mu_h^n \sigma_{hl}^n + \mu_l^n \sigma_{ll}^n)}, \text{ for } f = n, v.$$

In a similar vein, the average level of productivity over all startups in the ex post low state is given by

$$x_l = \frac{s^v(\mu_h^v \sigma_{hl}^v x_{hl}^v + \mu_l^v \sigma_{ll}^v x_{ll}^v) + s^n(\mu_h^n \sigma_{hl}^n x_{hl}^n + \mu_l^n \sigma_{ll}^n x_{ll}^n)}{s^v(\mu_h^v \sigma_{hl}^v + \mu_l^v \sigma_{ll}^v) + s^n(\mu_h^n \sigma_{hl}^n + \mu_l^n \sigma_{ll}^n)}, \tag{10}$$

with the fraction of all successful startups in this state being

$$\nu_l = \frac{s^v(\mu_h^v \sigma_{hl}^v x_{hl}^v + \mu_l^v \sigma_{ll}^v x_{ll}^v) + s^n(\mu_h^n \sigma_{hl}^n x_{hl}^n + \mu_l^n \sigma_{ll}^n x_{ll}^n)}{s^v(\mu_h^v \sigma_{hh}^v + \mu_l^v \sigma_{lh}^v) + s^n(\mu_h^n \sigma_{hh}^n + \mu_l^n \sigma_{lh}^n) + s^v(\mu_h^v \sigma_{hl}^v + \mu_l^v \sigma_{ll}^v) + s^n(\mu_h^n \sigma_{hl}^n + \mu_l^n \sigma_{ll}^n)}.$$

Next, define the following growth factors for the above three ex post classes of enterprises:

$$g_h^f \equiv \frac{x_h^f}{\mathbf{x}}, \text{ for } f = v, n, \text{ and } g_l \equiv \frac{x_l}{\mathbf{x}}.$$

The g's are determined by the choice problems (P1) and (P2) and must be constant along a balanced growth path. For this to be the case, the x_{ij}^f's have to grow over time at the same rate as $\mathbf{x}$. Additionally, the σ_{ij}^f's must be constant. Can a balanced growth path exist where entrepreneurs are choosing different levels of innovation? If so, what determines the aggregate growth rate for the baseline level of productivity, $g_{\mathbf{x}}$?

Now, suppose that the baseline level of productivity in the current period, $\mathbf{x}$, reflects the average productivities of each of the above three classes of enterprises as well as their mass in the set of producing startups. Specifically, let $\mathbf{x}'$ read

$$\mathbf{x}' = c(\nu_h^v x_h^v)^{\gamma_h^v}(\nu_h^n x_h^n)^{\gamma_h^n}(\nu_l x_l)^{1-\gamma_h^v-\gamma_h^n}, \tag{11}$$

where c is a constant term and $0 < \gamma_h^v,\ \gamma_h^n,\ \gamma_h^v + \gamma_h^n < 1$. This law of motion depends upon both the extensive (the ν's) and intensive (the x's) margins of the three ex post categories of startups. The aggregate level of productivity's gross growth rate, $g_{\mathbf{x}}$, will be

$$\begin{aligned} g_{\mathbf{x}} &\equiv \frac{\mathbf{x}'}{\mathbf{x}} = \frac{c(\nu_h^v x_h^v)^{\gamma_h^v}(\nu_h^n x_h^n)^{\gamma_h^n}(\nu_l x_l)^{1-\gamma_h^v-\gamma_h^n}}{\mathbf{x}} \\ &= c(\nu_h^v g_h^v)^{\gamma_h^v}(\nu_h^n g_h^n)^{\gamma_h^n}(\nu_l g_l)^{1-\gamma_h^v-\gamma_h^n}. \end{aligned} \tag{12}$$

because $g_h^v \equiv x_h^v/\mathbf{x}$, etc. Therefore, the gross growth rate in the baseline level of productivity, or $g_{\mathbf{x}}$, is just a geometric average of the average rates of innovation by each type of enterprise, or the g's, weighted by their respective masses, the ν's. The extent to which high-type VC-funded enterprises propel the economy forward is governed by g_h^f, γ_h^f, and ν_h^f, $f = \nu, n$.

Next, the market discount factor, δ, is given by

$$\delta = \beta / g_{\mathbf{x}}^{\rho\zeta/(\zeta+\lambda)}, \tag{13}$$

where β is the representative agent's subjective discount factor and ρ is the coefficient of relative aversion.[27] If the depreciation factor on capital is $\mathfrak{d}$, then the rental rate, r, will be given by

$$r = 1/\delta - \mathfrak{d}. \tag{14}$$

How is the wage rate determined in general equilibrium and at what gross rate will it grow along a balanced growth path? To address this question, first note that there are eight types of enterprises producing: viz., ex ante high- and low-type enterprises that turn into ex post high- and low-type firms, which were initially financed by venture capitalists and non-venture capitalists. Recall that each period there are s^v firms funded by VC and s^n by non-VC. Let $\#_{ij}^f$ refer to the number of current startups that will turn into type-(f, i, j) enterprises in the subsequent period. It is easy to deduce that

$$\#_{ij}^f = \sigma_{ij}^f s_f, \text{ for } f = v, n \text{ and } i, j = h, l.$$

Now, the survival rate for an operating firm is $\mathbf{s}$. Thus, the numbers of t-period-old type-(f, i, j) enterprises that will still be around are

$$\#_{ij,t}^f = \mathbf{s}^t \#_{ij}^f, \text{for } f = v, n;\ i, j = h, l; \text{ and } t = 0, 1, \cdots.$$

Here an age-0 enterprise is one that just starts producing in the current period.

[27] That is, in the background there is a representative consumer/worker who inelastically supplies one unit of labor and has a utility function (in period 1) of the form

$$\sum_{t=1}^{\infty} \beta^{t-1} c_t^{1-\rho}/(1-\rho),$$

where c_t is his consumption in period t.

Next, the generic demand for labor by a type-(f,i,j) enterprise, arising from the first-order condition for labor, is

$$l_{ij}^{f}=\left(\frac{\kappa}{r}\right)^{\kappa/\zeta}\left(\frac{\lambda}{w}\right)^{(\zeta+\lambda)/\zeta}x_{ij}^{f}z_{ij}^{f},\text{ for }f=v,n\text{ and }i,j=h,l.$$

Hence, the labor demanded by an enterprise that came on line t periods ago is

$$l_{ij,t}^{f}=\left(\frac{\kappa}{r}\right)^{\kappa/\zeta}\left(\frac{\lambda}{w}\right)^{(\zeta+\lambda)/\zeta}\frac{x_{ij}^{f}}{g_{\mathbf{x}}^{t+1}}z_{ij}^{f}=\frac{l_{ij}^{f}}{g_{\mathbf{x}}^{t+1}},\tag{15}$$

since $x_{ij,t}^{f}=x_{ij}^{f}/g_{\mathbf{x}}^{t+1}$. Thus, the total demand for labor by type-(f,i,j) enterprises will be

$$\sum_{t=0}^{\infty}\#_{ij,t}^{f}l_{ij,t}^{f}=\sum_{t=0}^{\infty}\mathbf{s}^{t}\#_{ij}^{f}\frac{l_{ij}^{f}}{g_{\mathbf{x}}^{t+1}}=\frac{\#_{ij}^{f}l_{ij}^{f}}{g_{\mathbf{x}}(1-\mathbf{s}/g_{\mathbf{x}})}.$$

The labor-market clearing condition will appear as

$$\sum_{f=v,n}\sum_{i=h,l}\sum_{j=h,l}\frac{\#_{ij}^{f}l_{ij}^{f}}{g_{\mathbf{x}}(1-\mathbf{s}/g_{\mathbf{x}})}=1,$$

Therefore, the solution for wages is given by

$$w=\lambda\left(\frac{\kappa}{r}\right)^{\kappa/(\zeta+\lambda)}\left[\frac{\sum_{f=v,n}\sum_{j=h,l}\sum_{j=h,l}\#_{ij}^{f}z_{ij}^{f}x_{ij}^{f}}{g_{\mathbf{x}}(1-\mathbf{s}/g_{\mathbf{x}})}\right]^{\zeta/(\zeta+\lambda)}\mathbf{x}^{\zeta/(\zeta+\lambda)}.\tag{16}$$

Focus on the right-hand side of the above equation. Along a balanced growth path, the $\#_{ij}^{f}$'s, x_{ij}^{f}'s, z_{ij}^{f}'s and $g_{\mathbf{x}}$ will all be constant. The only thing growing is $\mathbf{x}$. Therefore, wages grow at the gross rate

$$g_{w}=\left(\frac{\mathbf{x}'}{\mathbf{x}}\right)^{\zeta/(\zeta+\lambda)}=g_{\mathbf{x}}^{\zeta/(\zeta+\lambda)}.\tag{17}$$

The upshot of the above discussion is summarized by the lemma below.

Lemma 2 *(Balanced Growth) A balanced growth exists where*

1. *R&D by VC-backed startups, or σ_{ij}^{f} and x_{ij}^{f} for $i,j=h,l$, is determined by the Nash Bargaining problem (P1), taking as given δ, E_i, r, w, g_w, V, and $\mathbf{x}$.*

2. *R&D by non-VC-backed startups, or σ_{ij}^{n} and x_{ij}^{n} for $i,j=h,l$, and the threat point for talented entrepreneurs, E_i, are determined by the bank finance problem (P2), taking as given δ, i, r, w, g_w, and $\mathbf{x}$.*

3. *The wage rate, w, clears the labor market in accordance with (16), taking as given σ_{ij}^{f}, x_{ij}^{f}, and s_f, for $f = v, n$ and $i, j = h, l$, and $g_{\mathbf{x}}$. The gross growth rate in wages, g_w, is determined by (17) given $g_{\mathbf{x}}$.*

4. *The market discount factor, δ, and the rental rate on capital, r, are specified by (14) and (13), assuming a value for $g_{\mathbf{x}}$. The gross interest rate on bank loans, $1/i_i$, is determined by (7), conditional on values for δ, σ_{ih}^{n}, and σ_{il}^{n}.*

5. *The threat point for an entrepreneur, E_i, is determined by the outcome to problem (P2), assuming solutions for δ, i_i, w, g_w, and $\mathbf{x}$, for $i = h, l$, while the threat point for a venture capitalist is simply $V = 0$.*

6. *Aggregate productivity, $\mathbf{x}$, and its gross growth rate, $g_{\mathbf{x}}$, are specified by (11) and (12).*

From the above analysis of the balanced growth path, it is easy to deduce that a corporate income tax will have no impact on economic growth. This transpires because it acts as an equiproportionate reduction in all synergies, or on the z_{ij}^{f}'s. This operates in the same manner as a once-and-for-all reduction in aggregate productivity.

Lemma 3 *(Neutrality of a Uniform Corporate Income Tax on Long-Run Growth) Suppose the government levies a corporate income tax at rate τ on all producing enterprises, with all revenue is rebated back to consumers/workers in the form of lump-sum taxes. There is no effect on long-run growth; i.e., on $g_{\mathbf{x}}$ and g_w.*

Proof. Again, see Appendix 11. ■

6 Matching the Model with U.S. Census Data

The model is now confronted with stylized facts computed using data from the U.S. Census Bureau's LBD, the USPTO's PatentsView, and VentureXpert. The length of a period is taken to be 10 years. The baseline level of aggregate productivity, $\mathbf{x}$, can be innocuously normalized to one; hence, $\mathbf{x} = 1$. Some of the parameters in model are standard in macroeconomics. These are assigned common values. Parameters unique to the current study are calibrated in the recursive fashion discussed below. The calibration procedure uses facts for the average growth rate of the U.S. economy, the number of firms per worker in the economy, and the ratio of non-VC-funded to VC-funded startups. The ex ante matching probabilities, or the four μ_i^f's in Table 3, can be computed using data from the U.S. Census Bureau and VentureXpert. Likewise, the success odds for type-(f,i,j) ventures, or the eight σ_{ij}^f's, can be calculated using the same data. For each funding source, f, these amount to the transition probabilities of going from ex ante state i to ex post state j. Due to data restrictions imposed by the Bureau for reasons of confidentiality, facts can only be used for the average employment of type-(f,j) firms; that is, for the four observations $l_j^f = (\mu_h^f \sigma_{hj}^f l_{hj}^f + \mu_l^f \sigma_{lj}^f l_{lj}^f)/(\mu_h^f \sigma_{hj}^f + \mu_l^f \sigma_{lj}^f)$, for $f = v, n$ and $j = h, l$. (In other words, the data cannot also be categorized by the ex ante type i of enterprise.) Likewise, facts can only be used for the average patenting activity of type-(f,j) firms. These restrictions impose some limitations on the structure that can be calibrated. Given the limitations on the structure, there is exact identification for most of the parameter values unique to this study.

6.1 Number of Startups, s^n and s^v

From the U.S. data the success rates for the eight types of startups are known. That is, σ_{ij}^f can be assigned the value $\widetilde{\sigma}_{ij}^f$, for $f = v, n$ and $i, j = h, l$. A $\sim$ over a variable denotes that its value at the calibration point can be assigned from the data. Let π_j^f represent the fraction of all initial startups that were funded by source f and ended up in ex post state j. It is easy to deduce that

$$\pi_j^f = \mu_h^f \sigma_{hj}^f + \mu_l^f \sigma_{lj}^f < 1, \text{ for } f = v, n \text{ and } j = h, l. \tag{18}$$

The π_j^f's are known since the μ_i^f's and σ_{ij}^f's are. They lie below one because some startups will fail. The ex ante matching probabilities in the U.S. data are reported in Table 4. Venture capitalists fund a higher percentage of high-type firms than do other financiers.

Ex Ante Matching Probabilites (%)

Type	$\geq$ 90th	$\leq$ 90th
VC Funded	32.55	67.45
Non-VC Funded	10.64	89.36

Table 4: The data sources are discussed in Section 2.1. A firm's type refers to whether or not it is in the top decile for employment 10 years after VC funding.

In the United States the number of entrepreneurs (employer businesses) per worker is 4.3 percent. Therefore, it should transpire that

$$\frac{s^v(\widetilde{\pi}_h^v + \widetilde{\pi}_l^v) + s^n(\widetilde{\pi}_h^n + \widetilde{\pi}_l^n)}{(1-\mathbf{s})} = 0.043.$$

Additionally, the ratio of non-VC to VC-funded startups can be calculated from the U.S. data, so that a value for $\widetilde{s^n/s^v}$ is pinned down. Hence, using the above equation

$$s^v = 0.043 \times (1-\mathbf{s})/[(\widetilde{\pi}_h^v + \widetilde{\sigma}_l^v) + (\widetilde{s^n/s^v})(\widetilde{\pi}_h^n + \widetilde{\pi}_l^n)],$$

This equation can be used to back out values for s^v and $s^n = (\widetilde{s^n/s^v})s^v$.

6.2 Growth in Startup Productivities, the g_j^f's

The model is calibrated so that the benchmark model economy grows at 1.8 percent a year. This implies a target growth rate for aggregate productivity, $\widetilde{g}_{\mathbf{x}}$, denoted by

$$\widetilde{g}_{\mathbf{x}} = 1.018^{10\times(\zeta+\lambda)/\zeta}.$$

Data on patent stocks is used to identify the innovation rates of firms. As discussed, stylized facts for patenting activity can only be reported at the level of type-(f,j) firms; i.e., it cannot be broken down by the initial type, i, of the enterprise. So, the calibration procedure will impose the restriction that $x_{hj}^f = x_{lj}^f \equiv x_j^f$, for $f = v, n$ and $j = h, l$. The innovation rates for a type-(f,j) enterprise relative to a type-(v,h) firm are represented by $\widetilde{\phi}_j^f$. Then,

$$x_j^f = \widetilde{\phi}_j^f x_h^v, \text{ for } (f,j) \neq (v,h).$$

From equation (12) it then follows that

$$\widetilde{g}_{\mathbf{x}} = c x_h^v (\widetilde{\nu}_h^v)^{\gamma_h^v} (\widetilde{\nu}_h^n \widetilde{\phi}_h^n)^{\gamma_h^n} (\widetilde{\nu}_l \widetilde{\phi}_l)^{1-\gamma_h^v-\gamma_h^n},$$

Patent Citations (%)		
Type	$\geq$ 90th	$\leq$ 90th
VC Funded	26.22	2.67
Non-VC Funded	39.17	31.93

Table 5: Again, the data sources are discussed in Section 2.1. A firm's type refers to whether or not it is in the top decile for employment 10 years after VC funding.

where using (10) it is easy to deduce that

$$\phi_l \equiv \frac{s^v(\mu_h^v\sigma_{hl}^v + \mu_l^v\sigma_{ll}^v)\widetilde{\phi}_l^v + s^n(\mu_h^n\sigma_{hl}^n + \mu_l^n\sigma_{ll}^n)\widetilde{\phi}_l^n}{s^v(\mu_h^v\sigma_{hl}^v + \mu_l^v\sigma_{ll}^v) + s^n(\mu_h^n\sigma_{hl}^n + \mu_l^n\sigma_{ll}^n)}.$$

Now, choose $c = 1/[(\widetilde{\nu}_h^v)^{\gamma_l^v}(\widetilde{\nu}_h^n)^{\gamma_h^n}(\widetilde{\nu}_l)^{1-\gamma_h^v-\gamma_h^n}]$, an innocuous normalization. Then,

$$g_h^v = x_h^v = \frac{\widetilde{g}_{\mathbf{x}}}{(\widetilde{\phi}_h^n)^{\gamma_h^n}(\widetilde{\phi}_l)^{1-\gamma_h^v-\gamma_h^n}} \text{ (recall } \mathbf{x} = 1).$$

Therefore,

$$g_j^f = \widetilde{\phi}_j^f g_h^v, \text{ for } (f,j) \neq (v,h).$$

The spillover parameters, or the γ_j^f's, are computed from the shares of type-(f,j) firms in total patent citations. The distribution of patent citations by firm type in the United States is displayed in Table 5. On this account, $\gamma_h^v = 0.2622$, $\gamma_l^v = 0.0267$, and $\gamma_h^n = 0.3917$.

6.3 Synergies, the z_{ij}^f's

The relative employments of firms will be used to identify the synergy parameters or the z_{ij}^f's. Let $\widetilde{e}_j^f$ denote the employment of a type-(f,j) firm relative to a type-(v,h) one. There are four of these. As was discussed earlier, unfortunately employment cannot categorized by the ex ante type i of an enterprise. Given this only four unique synergy parameters can be backed out using U.S. Census data from the calibration procedure. So, set $z_{hj}^f = z_{lj}^f \equiv z_j^f$ in the calibration procedure. The four synergy parameters, z_h^v, z_l^v, z_h^n, and z_l^n, are easy to pin down. First, set $z_h^v = 1$. Then, using the labor demand functions (15) and the data on relative employments, $\widetilde{e}^f_j$, the following holds

$$\frac{x^v_h z^v_h}{x^f_j z^f_j} = (\frac{x^v_h}{x^f_j})\frac{1}{z^f_j} = \widetilde{e}^f_j,$$

so that

$$z^f_j = (\frac{x^v_h}{x^f_j})/\widetilde{e}^f_j, \text{ for } (f,j) \neq (v,h). \tag{19}$$

The wage rate, w, can be now determined at the calibration point using (16).

6.4 The Efficiencies of Research and Development, the ψ^f_{ij}'s, χ^f_{ij}'s, and ι

The desired values for x^f_{ij} and σ^f_{ij} can be achieved by backing out the constant parameters for the research and development functions, χ^f_{ij}'s and ψ^f_{ij}'s, so that the model hits these targets. In particular, the first-order condition for σ^f_{ij} states that

$$\delta x^f_{ij} z^f_{ij} I(\cdot') = w[\frac{1}{(1-\widetilde{\sigma}^f_{ij})^2} - (\frac{1}{1-\widetilde{\sigma}^f_{ij}} - 1)]/\chi^f_{ij},$$

so that

$$\psi^f_{ij} = \{w[\frac{1}{(1-\widetilde{\sigma}^f_{ij})^2} - (\frac{1}{1-\widetilde{\sigma}^f_{ij}} - 1)]\}/[\delta x^f_{ij} z^f_{ij} I(\cdot')]. \tag{20}$$

Following a similar procedure for x^f_{ij}'s gives

$$\chi^f_{ij} = w(\frac{x^f_{ij}}{\mathbf{x}})^{\iota-1}/[\mathbf{x}\delta\widetilde{\sigma}^f_{ij} z^f_{ij} \times I(\cdot')]. \tag{21}$$

Since there are eight values for the $\widetilde{\sigma}^f_{ij}$'s, there will be eight unique value for the χ^f_{ij}'s and ψ^f_{ij}'s (even though there are only four unique values for the x^f_{ij}'s).

A value also needs to chosen for ι, the elasticity parameter on the research cost function. Henrekson and Sanandaji (2018) report that Canada had a tax rate on VC-funded startups of 32 percent and a VC-investment-to-GDP ratio of 0.04 percent. Their numbers for the United States are 15 and 0.20 percent. This implies a tax elasticity of $-\ln(0.20/0.04)/\ln(0.15/0.32) = 2.12$. The model is fit to match this elasticity.

Calibrated Parameter Values

Parameter Values	*Description*	*Identification*
Consumers		
$\widehat{\delta} = 0.95$	Discount factor—annualized	7% risk-free rate
Firms		
$\kappa = 1/3 \times 0.85 = 0.283$	Capital's share	Standard
$\lambda = 2/3 \times 0.85 = 0.567$	Labor's share	Standard
$1 - \mathfrak{d} = 0.07$	Depreciation rate—annualized	Standard
$\mathbf{s} = 0.96$	Firm survival rate—annualized	Expected life of Compustat firms
Research		
$\chi_{hh}^{v} = 41.270, \chi_{hl}^{v} = 37.540$	VC Research efficiency, x_{ij}^{v}	VC pat./succ. rates, U.S. data
$\chi_{lh}^{v} = 57.49, \chi_{ll}^{v} = 23.610$		
$\chi_{hh}^{n} = 1.616, \chi_{hl}^{n} = 0.341$	Non-VC Research efficiency, x_{ij}^{n}	Non-VC pat./succ. rates, U.S. data
$\chi_{lh}^{n} = 18.780, \chi_{ll}^{n} = 0.193$		
$\iota = 3.5$	Research cost elasticity, x_{ij}^{v} and x_{ij}^{n}	Canada/U.S. Tax elasticity
Development		
$\psi_{hh}^{v} = 0.002, \psi_{hl}^{v} = 0.042$	VC Development efficiency, σ_{ij}^{v}	VC pat./succ. rates, U.S.data
$\psi_{lh}^{v} = 0.001, \psi_{ll}^{v} = 0.071$		
$\psi_{hh}^{n} = 0.006, \psi_{hl}^{n} = 0.249$	Non-VC Development efficiency, σ_{ij}^{n}	Non-VC pat./succ. rates, U.S. data
$\psi_{lh}^{n} = 0.000, \psi_{ll}^{n} = 0.597$		
Synergies		
$z_{h}^{v} = 1.00, z_{l}^{v} = 0.042$	VC	Empl. Ratios, U.S. data
$z_{h}^{n} = 0.801, z_{l}^{n} = 0.083$	Non-VC	
Ex ante Matching Odds		U.S. Data
$\mu_{h}^{v} = 0.326, \mu_{l}^{v} = 0.675$		
$\mu_{h}^{n} = 0.106, \mu_{l}^{n} = 0.894$		
Growth Externality		
$\gamma_{h}^{v} = 0.26, \gamma_{l}^{v} = 0.03$	Exponent on VC innovation	Citation distribution, U.S. Data
$\gamma_{h}^{n} = 0.39, \gamma_{l}^{n} = 0.32$	Exponent on non-VC innovation	
Model Period	10 years	VC contracts

Table 6: The parameter values used to replicate the data targets in Table 7.

6.5 The Model's Fit

The calibration targets for the United States and the model's fit for these targets are shown in Table 7, with the parameter values used being presented in Table 6. As can be seen, over the time period in question GDP grew at 1.8 percent a year. The model matches this number. A VC-funded startup is much more likely to successful than a non-VC-funded one. A VC-funded firm that starts out in the decile of top employment has a 33 percent chance of making it into the top employment decile 10 years later, as Table 7 shows. This compares with 23.77 percent of non-VC-funded firms. A VC-funded startup that begins in the bottom 90 percent of the employment distribution still has a 23.77 percent chance of making it to the top, compared with only 1.68 percent for a non-VC-funded firms. Recall from Table 4, though, that a VC-funded firm is much more likely to start out in the top decile than a non-VC-funded one; viz., 32.55 versus 10.64 percent. This is due to VC selection. The unconditional probabilities for VC-funded and non-VC-funded firms ending up in the high and low states are presented in Table 8. It can be seen that VC-funded firms have 26.81 and 9.06 percent chances of ending up in the high and low states, while for non-VC-funded startups the numbers are 3.58 and 31.71 percent. Conditioned on success, VC-funded firms have a much higher likelihood of winding up in the high state (74.75 percent) relative to non-funded ones (10.15 percent). The failure rates for both types of startups are very high, 64.1 and 64.7 percent.

A VC-funded startup that makes into the top 10 percent of firms has a much higher level of employment than a non-VC-funded firm. The latter's employment is only 16 percent of the former's. The employment levels of both types of startups, should they end up in the bottom 90 percent, are much smaller. Still, a VC-funded firm has 1.6 times the level of employment of a non-VC-funded firm. On average a VC-funded firm has 38.02 times [$\cong (0.75 \times 100 + 0.25 \times 0.59)/(0.10 \times 16.23 + 0.90 \times 0.36)$] the employment of a non-VC-funded one. Now, only a small number of startups are funded by venture capitalists. Specifically, there are 1,035.34 non-VC-funded startups for every VC-funded one. Given the superior performance of VC-funded firms, they account for 3.7 percent of employment

Calibration Targets

Target	*U.S. data*		*Model*	
Economic growth—annualized	1.80		1.80	
Success Rates, %	Top 10%	Bottom 90%	Top 10%	Bottom 90%
VC funded				
ex ante top 10%	33.11	6.48	33.11	6.48
ex ante bottom 90%	23.77	10.30	23.77	10.30
Non-VC funded				
ex ante top 10%	19.53	18.83	19.53	18.83
ex ante bottom 90%	1.68	33.24	1.68	33.24
Employment, Relative %	Top 10%	Bottom 90%	Top 10%	Bottom 90%
VC funded	100.00	0.59	100.00	0.59
Non-VC funded	16.23	0.36	16.23	0.36
Patenting, Relative %	Top 10%	Bottom 90%	Top 10%	Bottom 90%
VC funded	100.00	14.09	100.00	14.09
Non-VC funded	20.27	4.33	20.27	4.33
Firms per Worker, %	4.30		4.30	
Ratio non-VC to VC Startups	1,035.34		1,035.34	
Tax Elasticity	2.1		2.05	

Table 7: See Section 2.1 for the data sources. Average employment and patent stocks are measured with respect to VC-funded firms in the top 10 percent. Patenting is measured using a quality-adjusted patent stock. Transition probabilities (the success rates) give the likelihood that an ex ante type-i firm ends up as an ex post type-j firm 10 years after funding, for i,j = top 10 percent, bottom 90 percent.

Unconditional Success Probabilities (%)

	By state π_j^f		*By state*$\vert\sim$ failing $\pi_j^f/(\pi_h^f+\pi_l^f)$		*Overall* $\pi_h^f+\pi_l^f$	*Failure* $1-\pi_h^f+\pi_l^f$
Type	$\geq$ 90th	$\leq$ 90th	$\geq$ 90th	$\leq$ 90th		
VC Funded	26.81	9.06	74.75	25.25	35.87	64.13
Non-VC Funded	3.58	31.71	10.15	89.85	35.29	64.71

Table 8: Success probabilities when not conditioning on the initial type of startup, U.S. data and model. The figures are computed by using equation (18) and the information in Tables 4 and 7.

Rates of Innovation, Annualized %		
Rate of Innovation	*Top 10%*	*Bottom 90%*
VC funded	29.23	6.23
Non-VC funded	10.17	-5.59

Table 9: Annualized rates of innovation for each type of firm as measured by their growth from the frontier at the time of startup, $100 \times [(x_{ij}^{f}/\mathbf{x})^{1/10} - 1]$.

[$\cong 0.36 \times 38.02/(0.36 \times 38.02 + 1035.34 \times 0.35)$] of all firms 10 years after first funding. VC-funded firms however account for only 0.1 percent of startups. Hence, their punch far exceeds their weight.

Likewise, the patenting stock of a VC-funded startup in the top 10 percent of firms is much greater than that for a non-VC-funded firm. Non-VC-funded startups have only 20.27 percent of the patents that VC-funded startups have even when they resolve into the top decile of firms. VC-funded firms in the bottom 90th percentile of firms still have 3 times the patent stocks of non-VC-funded firms. VC-funded firms account for 1.2 percent of patenting activity, again even though they are just 0.1 percent of startups. Last, in the United States there are 23 workers (1/0.043) for every firm, which the model matches in addition to above statistics.

Table 9 shows that in the calibrated equilibrium most innovation in the economy is driven by VC-funded firms that make it into the top 10 percent. In fact, non-VC-funded startups that don't make it into the top decile of firms drag down the average level of productivity in the economy in the sense that even after they innovate they will lag behind the average level of productivity in the economy at the time of their inception. Additionally, they make up 89.77 percent of the firms in the economy, whereas VC-funded firms in the top decile comprise only 0.07 percent.

6.6 Interpretation of Backed-Out Parameter Values

The backed-out parameter values are reasonable, albeit there are anomalies. As can be seen from Table 6, the synergy level between an entrepreneur and the financier is highest for VC-funded firms in the ex post high state; i.e., $z_h^v > z_j^f$, for all $(f, j) \neq (v, h)$. Also, synergy levels are always higher in the high state as opposed to the low state, $z_h^f > z_l^f$, for all f. The synergy between entrepreneurs and financiers is very low in the ex post low state. Somewhat surprisingly, it is lowest for VC-funded firms; i.e., $z_l^v < z_l^n$. The reason is that non-VC-funded firms in the low state do almost no innovating; i.e., their productivity level, or x_l^n, is very small. Therefore, to match employment for these firms, the synergy level must be raised a little to compensate—in line with (19). Of course, VC-funded firms in the ex post low state are still better than non-VC-funded ones because $z_l^v x_l^v > z_l^n x_l^n$.

The development parameters also make some sense. First, it is more costly to achieve success in the ex post high state versus the low one. This is reflected by the fact that $\psi_{ih}^f < \psi_{il}^f$, for all f and i, where a higher value for ψ represents a lower cost. Second, if a firm starts out in the ex ante low state, as opposed to the high state, it is harder to end up in the ex post high state; i.e., $\psi_{lh}^f < \psi_{hh}^f$, for all f. Still, it is more expensive for a firm starting out as a high type to achieve success in the low state as opposed to one starting out as a low-type one; that is, $\psi_{hl}^f < \psi_{ll}^f$. This occurs because the observed odds of winding up in the low state are smaller for ex ante high-type firms. Hence, the calibration procedure places a high cost on this transition—see (20)—which is achieved by lowering ψ_{hl}^f relative to ψ_{ll}^f.

Last, the research parameters look reasonable, but again there are some exceptions. First, the cost of doing research is lower for a VC-funded firm than for a non-VC-funded one; that is, $\chi_{ij}^v \geq \chi_{ij}^n$ for all i, j. Thus, it is easy to do easier to do research when backed by a venture capitalist, perhaps because they lend expertise. Second, for a VC-backed firm the cost of researching an ex post high idea starting out from the ex ante low state is lower than for the ex ante high state; i.e., $\chi_{lh}^v > \chi_{hh}^v$. This is an artifact of the Census data restriction on reporting patenting activity that imposed the constraint $x_{lh}^v = x_{hh}^v$, when almost certainly $x_{lh}^v < x_{hh}^v$. Now, note that $z_{lh}^v = z_{hh}^v$ and that $\sigma_{lh}^v < \sigma_{hh}^v$. Given the low success rate when starting out from the ex ante low state it must transpire that, to get the same level of patenting as the ex ante high state, the research costs in the ex ante low state must be lower—as (21) dictates.

Ex Ante Matching Probabilities

	Perfectly Assortative				*Random*	
Type	$\geq$ 90th	$\leq$ 90th	$\geq$ 90th	$\leq$ 90th	$\geq$ 90th	$\leq$ 90th
VC Funded	$\mathfrak{v}/s^v$	$(s^v - \mathfrak{v})/s^v$	100.00	0	10	90
Non-VC Funded	$(.1 \times s - \mathfrak{v})/s^n$	$(s^n - .1 \times s + \mathfrak{v})/s^n$	9.91	90.09	10	90

Table 10: The ex ante matching probabilities when matching is perfectly assortative, left panel, and completely random, right panel. The numbers used in the simulation are reported as percentages.

7 Matching Thought Experiment

Just how important is ex ante matching between entrepreneurs and venture capitalists? To examine this question, suppose instead that there is perfect assortative matching between entrepreneurs and venture capitalists. To control the experiment, fix the supply of venture capitalists at s^v. So out of s startups, s^n will be funded by banks and s^v by venture capitalists. There will be $.1 \times s$ entrepreneurs in the top decile of projects. The situation is illustrated by the lefthand panel of Table 10. Now, the maximum number of matches between startups in the top 10 decile and venture capitalists is $\mathfrak{v} \equiv \min\{s^v, .1 \times s\}$, where $s = s^n + s^v$. Therefore, the probability of a venture capitalist funding a startup in the top decile is $\mathfrak{v}/s^v$. When venture capitalists are in short supply, as in the simulation, these odds would be 100 percent; here $\mathfrak{v} = s^v$ matches will occur between entrepreneurs in the top decile and venture capitalists. In general, there will be $s^v - \mathfrak{v}$ venture capitalists left over. They will now have to fund firms below the top decile. The odds of such a match happening are given by $(s^v - \mathfrak{v})/s^v$. In the simulation, the pool of venture capitalists is exhausted by entrepreneurs in the top decile (because $s^v = \mathfrak{v}$). Therefore, this situation doesn't happen. Returning to the general case, there could potentially be $.1 \times s - v$ high-quality entrepreneurs that are left unfunded by venture capitalists. Thus, the probability of a match between a high-quality entrepreneur and a bank is $(.1 \times s - \mathfrak{v})/s^n$. Since s^n projects will be funded by banks, they can fund $s^n - .1 \times s + \mathfrak{v}$ startups lying outside of the top decile. Therefore, the probability of a bank funding a low-quality entrepreneur is $(s^n - .1 \times s + \mathfrak{v})/s^n$.

Matching Thought Experiment

	Baseline	*Perfectly Assortative*	*Random*
Variable			
Growth	1.8	1.96	1.78
VC firms, high			
fraction, ν_h^v	0.73	0.91	0.68
patenting, x_{hh}^v and x_{lh}^v	12.99, 12.99	13.13, 13.15	13.06, 13.07
success, σ_{hh}^v and σ_{lh}^v	33.11, 23.77	33.56, 24.18	33.12, 23.78
VC share of all patenting	1.28	1.5	1.20
Non-VC firms, high			
fraction, ν_h^n	99.28	99.09	99.32
patenting, x_{hh}^v and x_{lh}^v	2.63, 2.63	2.67, 2.68	2.55, 2.57
success, σ_{hh}^n and σ_{lh}^n	19.53, 1.68	19.91, 1.73	19.72, 1.71
Non-VC share of all patenting	98.72	98.5	98.8

Table 11: The experiments report what would happen if the initial matching between entrepreneurs and financiers is either perfectly assortative or random. All rates are reported as percentages.

The movement to perfect assortative matching increases growth from 1.8 to 1.96 percent—see Table 11. Virtually all of the uptick in growth derives from a movement along the extensive margin. That is, from an increase in the fraction of VC-funded firms that are operating in the ex post high state. This fraction moves up from 0.73 to 0.91 percent. This is partially offset by a decline in the fraction of non-VC-funded firms producing in this state. The percentage increase in VC-funded firms is however much larger. Effectively the number of non-VC funded firms barely moves from its baseline. Observe that the success and patenting rates increase a little. This occurs because in the momentary equilibrium the wage rate drops slightly. So, the increase in growth comes from an increase in the fraction of firms in the high state that are funded by venture capitalists. As can be seen, the contribution of VC firms to patenting activity rises from 1.28 to 1.55 percent. This has a noticeable impact on growth due to the sizeable spillover coefficient, γ_h^v, on VC patenting activity in the high state.

Alternatively, suppose that matching is random. The situation is given by the right-hand panel of Table 11. There are s^v venture capitalists in total. With random matching, $.1 \times s^v$ venture capitalists will be matched with firms in the top 10 percent, the rest with firms in the bottom 90 percent. Growth now drops, but ever so slightly—see Table 11. This occurs primarily from a fall in the number of VC-funded firms operating in the high state from 0.73 to 0.68 percent. The movement from the matching structure observed in the data to the random matching structure is much smaller than moving from the observed matching structure to perfect assortative matching, as can be seen from Tables 4 and 10. So, the impact on growth from random matching is more muted than the impact on growth from perfect assortative matching.

8 The Taxation of Startups

In the United States VC startups are taxed when they are floated or sold at the capital gains rate, currently 17 percent. This happens because the key officers of the companies are remunerated in the form of stock options and/or preferred shares that only attain value when the firm goes public or is bought by another company. Other forms of startups are taxed at either the corporate or personal income tax rates, which are much higher. Until recently the corporate income tax rate was 35 percent. Lemma 3 states that the taxation of all startups at a uniform rate would have no impact on an economy's balanced growth path. Differential taxation across VC-funded and non-VC funded startups will though. As an experiment, consider what would have happened if the VC- and non-VC-funded startups were taxed at a uniform rate of 35 percent. Note that the historical tax rates on VC- and non-VC-funded startups are built into the calibrated synergy factors, or the z's. So, the experiment amounts to degrading the synergy factor on VC-funded startups by a factor of $(1.0 - 0.35)/(1.0 - 0.17)$.[28]

When VC-funded startups are taxed at a higher rate, economic growth drops from 1.8 to 1.48 percent per year, as Table 12 shows. Not surprisingly, the level of patenting activity by VC-funded firms in the ex post high state drops. This is not made up for by an increase in patenting activity by non-VC-funded firms. Additionally, VC-funded firms' success rates for making it into the high state drops off sharply. The success rates for non-VC-funded firms does not rise to offset this. The share of VC firms in patenting activity drops fairly

[28] This can be gleaned by examining the proof for Lemma 3 in Appendix 11.

Uniform Taxation Experiment

Variable	*Baseline*	*Uniform Taxation*
Growth	1.8	1.48
VC firms, high		
fraction, ν_h^v	0.73	0.53
patenting, x_{hh}^v and x_{lh}^v	12.99, 12.99	10.44, 10.14
success, σ_{hh}^v and σ_{lh}^v	33.11, 23.77	24.79, 16.52
VC share of all patenting	1.28	0.74
Non-VC firms, high		
fraction, ν_h^n	99.28	99.47
patenting, x_{hh}^v and x_{lh}^v	2.63, 2.63	2.60, 2.59
success, σ_{hh}^n and σ_{lh}^n	19.53, 1.68	19.18, 1.64
Non-VC share of all patenting	98.72	99.26

Table 12: The experiment reports what happens when the tax rate on VC-funded startups is raised from 17 percent to the 35 percent rate on non-VC-funded startups. All rates are reported as percentages.

dramatically. Since VC-funded are the prime movers in the economy, all of this results in a decline in the growth rate.

9 Conclusion

An endogenous growth model is developed and then simulated where entrepreneurs take ideas for projects to either banks or venture capitalists for financing. In the U.S. data, the portfolio of startups financed by venture capitalist is much more promising in terms of growth and innovation, in an ex ante sense, than the portfolio of firms financed by banks. The initial selection process observed in the data is imposed on model. Additionally, in the U.S. data, firms that are financed by venture capitalists, as opposed to banks, have a much higher likelihood of making it into the top decile of firms in terms of employment ten years later. They also have much higher rates of growth and patenting activity, conditional on surviving. The model's saturated structure allows the framework to be calibrated so that it matches exactly the observed transitions from ex ante promise to ex post success for both bank- and VC-financed firms. It also matches the small ratio of VC-financed to bank-financed startups, the small average firm size in the U.S. economy, and GDP growth per capita. It is shown that the degree of ex ante assortative matching between startups and financiers matters for growth. Last, the differential taxation of bank- and VC-financed startups is important for growth.

References

Autor, David, David Dorn, Lawrence F. Katz, Christina Patterson, and John Van Reenen. 2017a. "The Fall of the Labor Share and the Rise of Superstar Firms." National Bureau of Economic Research Working Paper 23396.

Autor, David, David Dorn, Lawrence F. Katz, Christina Patterson, and John Van Reenen. 2017b. "Concentrating on the Fall of the Labor Share." *American Economic Review* 107(5):180-85.

Bayard, Kimberly, Dinlersoz, Emin, Dunne, Timothy, Haltiwanger, John, Miranda, Javier, and John Stevens. 2018. "Early-Stage Business Formation: An Analysis of Applications for Employer Identification Numbers." National Bureau of Economic Research Working Paper 24364.

Bernstein, Shai, Giroud, Xavier, and Townsend, Richard. 2014. "The Impact of Venture Capital Monitoring." *Journal of Finance* 71(4): 1591-1622.

Brown, David, Earle, John, Kim, Mee Jung, Lee, Kyung Min. 2018. "High-Growth Entrepreneurship.", IZA Discussion Paper 11662.

Catalini, Christian, Guzman, Jorge, and Scott Stern. 2017. "In Plain Sight: Venture Growth with or without Venture Capital." Unpublished paper, Dartmouth University.

Choi, Joonkyu. 2017. "Entrepreneurial Risk Taking, Young Firm Dynamics and Aggregate Implications." Unpublished paper, University of Maryland.

Decker, Ryan, John Haltiwanger, Ron Jarmin, and Javier Miranda. 2016. "Where Has All the Skewness Gone? The Decline in High-Growth (Young) Firms in the U.S." *European Economic Review* 86(2016): 4-23.

Gabaix, Xavier. 2011. "The Granular Origins of Aggregate Fluctuations." *Econometrica* 79(3): 733–772.

Graham, Stuart, Cheryl Grim, Tariq Islam, Alan Marco, and Javier Miranda. 2015. "Business Dynamics of Innovating Firms: Linking U.S. Patent Data with Administrative Data on Workers and Firms." Center for Economic Studies Working paper CES-WP-15-19.

Greenwood, Jeremy, Pengfei Han, and Juan M. Sanchez. 2018. "Financing Ventures." National Bureau of Economic Research Working Paper 24808.

Henrekson, Magnus and Tino Sanandaji. 2018. "Stock Option Taxation and Venture Capital Activity: A Cross-Country Study." *Venture Capital: An International Journal of Entrepreneurial Finance*. 20(1): 51-71.

Jovanovic, Boyan and Balazs Szentes. 2013. "On the Market for Venture Capital." *Journal of Political Economy*, 121 (3): 493-527.

Opp, Christian. 2018. "Venture Capital and the Macroeconomy." *Journal of Finance*, forthcoming.

Pugsley, Benjamin, Sedlacek, Petr, and Vincent Sterk. 2017. "The Nature of Firm Growth." Centre for Macroeconomics Discussion Paper 2017-37.

Puri, Manju and Rebecca Zarutskie. 2012. "On the Life Cycle Dynamics of Venture-Capital and Non-Venture-Capital-Financed Firms." *Journal of Finance* 67(6): 2247-2293.

Silveira, Rafael and Randall D. Wright. 2016. "Venture Capital: A Model of Search and Bargaining." *Review of Economic Dynamics*, 19: 232–246.

10 Appendix: Data

This Data Appendix describes the assembling of the main dataset and the analysis sample, and provides further detail on how tables and figures are constructed.

10.1 Data merging procedure

To construct the analysis data, the Thomson-Reuter's VentureXpert data on firms receiving VC is combined with the U.S. Census Bureau's Business Register (BR) using probabilistic name and address matching. The resulting data is then linked with the U.S. Census Bureau's Longitudinal Business Database (LBD).

The procedure to combine VentureXpert and BR involves several steps. The first step is to clean the two datasets to facilitate matching. For VC-backed firms between 1970 and 2015, VentureXpert provides the following information: firm identifier, firm name, street address, city, state, zipcode, and the year in which VC funding is received. The analysis uses only the firms located in the United States and identifies the first and last years of funding for all such firms. The BR is the main sampling frame for U.S. Census Bureau's surveys and it is the main source data for the LBD. Establishment and firm identifiers, years of operation, and employment from the LBD are integrated with the information on firm name, street address, city, state, and zip code from the BR for the period 1976 through 2013. The firm name, street address, city, and zip code variables in both the BR and VentureXpert are standardized using the same protocol.

The second step involves merging VentureXpert with the BR-LBD data in a sequential manner using nine passes. In each pass, firms are linked based on certain firm name and address criteria, and removed from the set of linked firms before proceeding to the next pass. In the first pass, the merge is based on firm name, street address, city, and state. In the second pass, the merge is performed using firm name, city, state, and zip code. In the third through fifth passes, the merge is done on firm name and two additional address criteria. In the sixth through eight passes, the merge utilizes firm name and one additional address criterion. Finally, in the ninth pass, only firm name is used for matching.

The implementation of the procedure described above often results in multiple BR-LBD links for each VC-funded firm. In the third stage, the best BR-LBD match is identified for each VC-funded firm using a composite match quality score, which is constructed based on the quality of the name, city, state, and zip code match. To determine the quality of a match, the Jaro-Winkler string comparator is utilized. For each VC-funded firm, only the highest quality BR-LBD links are kept based on the composite score. In the process, low-quality links are also dropped—these are the links made only based on name with a low Jaro-Winkler score, and only based on name and one address variable with a low Jaro-Winkler score.

After implementing these steps, there still remain multiple BR-LBD links for each VC-funded firm. The fourth stage involves validating these links using information on the years in which firms were funded. A link is considered valid if a firm is determined to be active within a five year window of the first and last year of VC funding. In particular, it is required that the funded firm appears in the LBD at least within the first five years of receiving its first VC funding, and that the funded firm not disappear (exit) from the LBD more than five years before receiving its last VC funding. Applying this window also helps account for some potential lags in reporting or recording errors. In cases where multiple links remain, firms whose first year of operation is closest to the founded year recorded by VentureXpert are kept. Remaining cases with multiple links are dropped to reduce false positives. Overall, the entire procedure results in a BR-LBD match rate of 70 percent for the VC-funded firms in VentureXpert.

10.2 Analysis sample construction

The empirical analysis focuses on VC-funded firms that receive funding between 1980 and 2012. This is the period with the best overlapping coverage across all three data sets—LBD, patenting data, and VentureXpert. The last year of the analysis is set to 2012 to take into account lags in the patenting data.

Each VC-funded firm is effectively matched to an establishment in the LBD, since the majority of firms in the United States are single-unit (single-establishment) firms, particularly when they are in the age range when most firms receive VC funding (i.e., when they are very young). As a result, for the VC-funded firm sample the focus is on firms that are born as single-unit firms between 1980 and 2012.

The establishment identifier is time invariant in the LBD across all the firms an establishment is associated with throughout its life. Hence, an establishment's expansion can be tracked, including the opening of additional establishments under the same firm and M&A (the funded establishment being subsumed into an existing firm). If the establishment identifier disappears, but the firm identifier associated with it in the last year of its existence continues, the latest firm the establishment belongs to is followed.

The empirical analysis compares the performance of VC-funded firms with all other non-VC-funded firms in the economy. For this purpose, a sample of non-VC-funded firms is assembled using a similar procedure. In particular, between 1980 and 2012 the analysis considers all non-VC-funded firms that are born as single-units and follows the associated establishment identifier across all the firms it is part of throughout its life. Firms that are born as multi-units (a very small fraction of new-born firms) are excluded, as well as those single-unit firms that have an initial age different than zero—these firms are parts of existing businesses.

For both the VC-funded and non-VC-funded firms in the sample, there is information on firm-level industry, age, location (state), multi-unit status, employment, and payroll. Years associated with significant events in the firm's life are also identified. These events are firm failure, merger/acquisition, and IPO. Firm failure is identified when all of the establishments of a firm cease to report employment and stop appearing in the LBD. Mergers and acquisitions are identified when one or more of the establishments belonging to a particular firm appear in subsequent years to be part of a different firm that existed in the two years prior to the acquisition, continues to exist for at least two years after the acquisition, and is a multi-unit firm, which means that it operates more than one establishment. IPOs are identified by the first year in which a firm meets two criteria: it appears in the Compustat-BR bridge constructed at the U.S. Census Bureau and records a positive stock price in Compustat.

For each VC-funded firm, the first year in which it received funding is also identified. In addition, firm-level patenting information is brought in using the LBD-PatentsView link file provided by Nikolas Zolas (U.S. Census Bureau). The links identify which firms in the sample engage in patenting activity and allow tracking of a number of innovations outcomes: the number of new patents, the stock of patents, the number of citations received by new patents, and the stock of citations based on patent application year. Two adjustments are made to the citation data. Self-citations are dropped and all citations are scaled by the average citations received by other patents applied for in the same year and same broad NBER technology class.

In the data when venture backed firms are acquired by existing firms, the firm identifier they are associated with changes. This change often leads to jumps in employment and patenting arising from the fact that the acquiring firm usually operates at a larger scale than the acquired firm. These abrupt changes in firm outcome measures can confound the analysis. Therefore, in cases where mergers or acquisitions are identified, the employment and patenting of the acquired firm are no longer tracked, and the last observed values are used.

10.3 Construction of figures and tables

10.3.1 Tables 1 and 15

In Tables 1 and 15, the outcomes of VC-funded and non-VC-funded firms are compared at the time of VC funding, as well as 10 years after VC funding. First, using the sample of

VC-funded firms, average firm age at first funding event is calculated. This age is used as the synthetic date to impute a first funding year for all non-VC-funded firms. Second, the analysis is restricted to firms for which the first actual or synthetic funding occurs in 2002 or earlier to ensure that survivors (firms that do not fail or get involved in M&A) are observed for a minimum of 10 years. For Table 1, employment at the time of first VC funding is used to separate firms into those in the top 10th percentile versus the bottom 90th percentile of firm employment at the time of first VC funding. These two groups are further divided into those firms that are funded by venture capitalists versus not. For each group their shares, at the time of first VC funding, in the number of firms and employment are calculated. Their shares in VC-funded versus non-VC-funded firms are also computed, as well as their average employment relative to VC-funded firms in the top 10th percentile.

In Table 15, the focus is on the subsample of firms that survive 10 years either after first actual or synthetic VC funding. If a firm is involved in an M&A activity before the 10 year mark, it is also included in the subsample. Firms are separated into two groups based on the distribution of employment 10 years after VC funding—those that fall into the top 10th percentile of the distribution versus the bottom 90th percentile. The two groups are then further divided into those that were funded by venture capitalists versus those that were not. As in Table 1, for each group statistics on their share of the number of firms and employment 10 years after first VC funding are calculated, as well as average employment and the patent stock relative to VC-funded firms in the top 10th percentile.

Transition probabilities between groups in Table 1 and Table 15 are also computed. These are the conditional probabilities with which a firm ends up in each of the four groups 10 years post first funding (Table 15), given which of the four groups it started in at the time of first funding (Table 1)—provided that such a transition is feasible.

10.3.2 Figure 2

To generate Figure 2, firm age at first VC funding is calculated as the difference between the first year in which the firm is observed as an employer business in the LBD and the first year of VC funding. The figure is restricted to those firms that receive funding between the ages of -5 and 10, where negative numbers indicate that funding was received prior to the first year in which the firm hired employees. Firms outside of this age range are not included to prevent disclosure of confidential information—they account for a small fraction of all firms that receive VC funding.

10.3.3 Figures 3 and 4

To generate Figure 3, the average Davis-Haltiwanger-Schuh (DHS) average annual employment growth rate in the firm's first three years of life is calculated. Firm births are associated with an annual growth rate of +2 in the first year of their lives. Firm failures are associated with a DHS growth rate of −2 in the year of failure. The firms in the analysis sample are then grouped into quintiles based on the distribution of DHS average employment growth rate, with the top quintile representing the highest average growth rates. For each quintile, the number of firms that subsequently receive VC funding is then divided by the total number of firms in that quintile to obtain the percentage plotted in the y-axis.

To generate Figure 4, the sample is restricted only to patenting firms, and the total quality-adjusted patenting of each firm is calculated for the first three years of patenting. Firms are then grouped into quintiles based on the quality-adjusted patenting measure, with the highest quintile corresponding to the largest amount of patenting. The y-axis is the ratio of the number of firms that subsequently receive VC funding to the total number of firms in that quintile (expressed as a percentage).

10.3.4 Figures 5, 6, and 7

In Figures 5, 6, and 7, the outcomes of funded firms are compared with those of the non-VC-funded firms. The employment, age, industry, and location distributions of VC-funded and non-VC-funded firms are quite different, which may lead to an overstatement of the differences between the two groups. To address these differences that indicate selection, for each VC-funded firm in the sample matching firms are found from the sample of non-VC-funded firms that are observationally similar to the VC-funded firm in the year first VC funding is received. Specifically, for each year 1980 through 2012 a firm that received first VC funding in a year is matched with a non-VC-funded firm in the same year that is of the same age, is in the same employment bin, has the same multi-unit status, is located in the same state, and operates in the same 4-digit NAICS industry.

Ultimately, the matching procedure assigns multiple matches to each VC-funded firm, each of which are given an equal weight in the analysis.[29] Of the nearly $19,000$ VC-funded firms in the full sample, approximately $13,000$ are matched with non-VC-funded firms that form the control group. The control group consists of a total of about $649,000$ firms. For each VC-funded firm, this procedure sometimes generates multiple matches from the non-VC-funded firm sample, owing to the large set of potential matches available. For the cases with multiple matches, only those matches are kept where the percent difference in initial employment (employment at age zero) between the VC-funded and non-VC-funded firm does not exceed 50 percent. After this elimination of multiple matches, there are still cases where multiple non-VC-funded firms match to a single VC-funded firm. In these cases, equal weights are assigned to each matching firm. In certain cases, one non-VC-funded firm also matches to multiple VC-funded firms. To ensure that a single trajectory is followed for the non-VC-funded firm, its earliest match is kept.

For the analysis of firm innovation, two patent-related outcome variables are considered. A separate set of matched VC-funded and non-VC-funded samples are created following the same procedure described above, but with the additional restriction that non-VC-funded matches have to be patenting firms—a restriction fulfilled by a relatively small number of firms.[30] The subsample of VC-funded firms that have patents consists of nearly $6,500$ firms. Of these about $2,500$ are successfully matched based on the criteria described above to about $5,000$ non-VC-funded firms that also have patents.

In Figure 5 the evolution of average employment for VC-funded and non-VC-funded firms is plotted. To construct the figure, the sample is restricted to those VC-funded firms (and their non-VC-funded matches) that received funding at ages 1, 2, or 3 and have non-zero employment. These restrictions ensure that the focus is on a sample where firm employment

[29] Out of the many matches for a firm, only those where the percentage difference between the VC-funded and non-VC-funded firm for both employment and initial employment (employment at age zero) do not exceed 50 percent are kept. Even after this filtering, in some cases multiple non-VC-funded firms match to a single VC-funded firm. In these cases, uniform weights are assigned for each match. For example, if 5 non-VC-funded firms match to a single VC-funded firm, each receives a weight of 0.2.

[30] No matching is done on patent quality because the set of potential control firms is very small.

is observed in at least one year prior to the first year of funding. In Figure 7 the sample is also restricted to those VC-funded firms (and their non-VC-funded counterparts) that receive their first funding at age 1, 2, or 3, but no additional restrictions are placed based on employment.

Using the matched sample, in Figure 6 the focus is on those firms that survive for at least 10 years post VC funding. For each firm, the DHS employment growth rate is calculated between the year in which funding is received and 10 years later. For both the VC-funded and non-VC-funded groups, the kernel density estimate of the resulting growth rate is plotted.

10.3.5 Figures 8 and 9, and Table 2

In Figures 8 and 9, and in Table 2, the focus is on high- versus low-quality VC funders. There are several steps involved in assigning VC-funded firms a funder quality measure. The VentureXpert data predominately consists of one observation per VC-funded firm per year. Each row of data has a list of funders and the average equity invested per funder. First, a separate indicator is created for each funder, which is used to generate a full list of VC funders in the data. Since the underlying VentureXpert data are obtained from regulatory disclosures, slight variations are observed in the spelling of funder names across deals and over time. In order to take these variations into account, alternate spellings of funder names in the data are taken into account to group these slight variations in spelling under the same identifier. These identifiers are then used to create a measure of total number of deals in which the funder was involved in over the years. A funder is defined to be of high quality (or high experience) if it is in the top 10th percentile of the total number of deals distribution, and of low quality (low experience) if it is in the bottom 90th percentile.

It is common for a VC-funded firm to receive funding from multiple venture capitalists within the first year of funding. If 50 percent or more of the firms' funders (in counts) are high quality, it is considered to be a firm funded by high quality venture capitalists. If, instead, more than 50 percent of its funders are low quality, it is considered to be a firm funded by low-quality venture capitalists. As a result, each VC-funded firm is unambiguously classified to either be funded by a high quality or a low-quality funder the first year it received VC funding.

In Figures 8 and 9 the sample is restricted to those VC-funded firms that receive their first funding at age 1, 2, or 3. The evolution of employment (Figure 8) and patenting (Figure 9) are separately plotted for those firms identified as being funded by high- versus low-quality venture capitalists. In Table 2, the outcomes of firms funded by high- and low-quality funders are compared more formally by using regression analysis. A VC-funded firm's employment 10 years after first VC-funding is used as a dependent variable. The dependent variables are an indicator of whether funding was provided by a high- or low-quality funder (omitted category is the low quality), age at which first funding is received, years since first funding event, initial firm size (employment) at funding, the amount of equity received at time of first funding, and 4-digit NAICS industry-year fixed effects.

11 Appendix: Theory

Proof. (The Impact of Synergy) The six first-order conditions connected with the Nash Bargaining problem (P1) are

$$\begin{aligned}&[\delta x_{ij}^{v} z_{ij} I(\cdot') - \delta p_{ij}^{v}] \times \{\sigma_{ih}^{v}\delta p_{ih} + \sigma_{il}^{v}\delta p_{il}^{v} - R(\frac{x_{ih}^{v}}{\mathbf{x}}) - R(\frac{x_{il}^{v}}{\mathbf{x}}) - D(\sigma_{ih}^{v}) - D(\sigma_{il}^{v}) - V\} \\ &\quad + \{\sigma_{ih}^{v}\delta x_{ih}^{v} z_{ih}^{v} I(\cdot') + \sigma_{il}^{v}\delta x_{il}^{v} z_{il}^{v} I(\cdot') - \sigma_{ih}^{v}\delta p_{ih}^{v} - \sigma_{il}^{v}\delta p_{il}^{v} - E_i\} \times [\delta p_{ij}^{v} - D_1(\sigma_{ij}^{v})] = 0,\end{aligned}$$

$$\begin{aligned}&\sigma_{ij}^{v}\delta z_{ij}^{v} I(\cdot') \times \{\sigma_{ih}^{v}\delta p_{ih} + \sigma_{il}^{v}\delta p_{il}^{v} - R(\frac{x_{ih}^{v}}{\mathbf{x}}) - R(\frac{x_{il}^{v}}{\mathbf{x}}) - D(\sigma_{ih}^{v}) - D(\sigma_{il}^{v}) - V\} \\ &\quad - \{\sigma_{ih}^{v}\delta x_{ih}^{v} z_{ih}^{v} I(\cdot') + \sigma_{il}^{v}\delta x_{il}^{v} z_{il}^{v} I(\cdot') - \sigma_{ih}^{v}\delta p_{ih}^{v} - \sigma_{il}^{v}\delta p_{il}^{v} - E_i\} \times R_1(\frac{x_{ij}^{v}}{\mathbf{x}}),\end{aligned}$$

and

$$\begin{aligned}&-\sigma_{ij}^{v}\delta \times [\sigma_{ih}^{v}\delta p_{ih} + \sigma_{il}^{v}\delta p_{il} - R(\frac{x_{ih}^{v}}{\mathbf{x}}) - R(\frac{x_{il}^{v}}{\mathbf{x}}) - D(\sigma_{ih}^{v}) - D(\sigma_{il}^{v}) - V] \\ &\quad + [\sigma_{ih}^{v}\delta x_{ih}^{v} z_{ih}^{v} I(\cdot') + \sigma_{il}^{v}\delta x_{il}^{v} z_{il}^{v} I(\cdot') - \sigma_{ih}^{v}\delta p_{ih}^{v} - \sigma_{il}^{v}\delta p_{il}^{v} - E_i] \times \sigma_{ij}^{v}\delta = 0,\end{aligned}$$

for $j = 1, 2$. Using the last first-order condition to cancel out the terms in braces in the first two statements leads to the two efficiency conditions (4) and (5). Displacing (4) and (5) with respect to z_{ij}^{v} yields

$$\begin{bmatrix} \delta z_{ij}^{v} I' & -D_{11} \\ -R_{11}/\mathbf{x}^2 & \delta z_{ij}^{v} I' \end{bmatrix} \begin{bmatrix} dx_{ij}^{v} \\ d\sigma_{ij}^{v} \end{bmatrix} = -\begin{bmatrix} \delta x_{ij}^{v} I' \\ \delta \sigma_{ij}^{v} I' \end{bmatrix} dz_{ij}^{v},$$

so that

$$\frac{dx_{ij}^v}{dz_{ij}^v} = \frac{-(\delta I')^2 x_{ij}^v z_{ij}^v - \delta\sigma_{ij}^v I' D_{11}}{\Delta} > 0,$$

and

$$\frac{d\sigma_{ij}^v}{dz_{ij}^v} = \frac{-(\delta I')^2 z_{ij}^v \sigma_{ij}^v - (R_{11}/\mathbf{x}^2)\delta x_{ij}^v I'}{\Delta} > 0,$$

with

$$\Delta \equiv (\delta z_{ij}^v I')^2 - D_{11}R_{11}/\mathbf{x}^2 < 0,$$

where the sign of Δ follows from the second-order condition for a maximum.[31] ■

Proof. (Neutrality of a Uniform Corporate Income Tax on Long-Run Growth) The proof is simple. The profits for an ex ante type-(f, i, j) enterprise will now be $(1-\tau)x_{ij}^f z_{ij}^f(1-\kappa-\lambda)[(\frac{\kappa}{r})^\kappa(\frac{\lambda}{w})^\lambda]^{1/\zeta}$. One could rewrite this as $x_{ij}^f \widetilde{z}_{ij}^f(1-\kappa-\lambda)[(\frac{\kappa}{r})^\kappa(\frac{\lambda}{w})^\lambda]^{1/\zeta}$, where $\widetilde{z}_{ij}^f \equiv (1-\tau)z_{ij}^f$. Hence, the introduction of a corporate income tax amounts to an equiproportional renormalization of all of the synergies. From equation (16), it can seen that wages will fall by a factor of $(1-\tau)^{\zeta/(\zeta+\lambda)}$. This fact, in conjunction with (3), implies that the value of an IPO will drop by this factor too. Last, as can be seen from the first-order conditions for σ_{ij}^f and x_{ij}^f, or (4), (5), (8), and (9), the marginal benefits and costs of research and development fall in lock step by same factor $(1-\tau)^{\zeta/(\zeta+\lambda)}$. Therefore, the solutions for σ_{ij}^f and x_{ij}^f are unchanged. ■

[31] The maximization problem linked to the above two efficiency condition is

$$\max_{\sigma_{ij}^v, x_{ij}^v} \{\sigma_{ij}^v \delta x_{ij}^v z_{ij}^v I(\cdot') - D(\sigma_{ij}^v) - R(x_{ij}^v/\mathbf{x})\}.$$

For the second-order condition for a maximum to hold, it must transpire that

$$\begin{vmatrix} -D_{11} & \delta z_{ij}^v I' \\ \delta z_{ij}^v I' & -R_{11}/\mathbf{x}^2 \end{vmatrix} > 0,$$

or that

$$D_{11}R_{11}/\mathbf{x}^2 - (\delta z_{ij}^v I')^2 > 0.$$