Dynamics of High-Growth Young Firms and the Role of Venture Capitalists[*]

Yoshiki Ando[†]

August 24, 2024 (Click here for the most recent version)

Abstract

How does venture capital (VC) financing help the growth of startup firms and impact aggregate output and consumption? Motivated by the substantial growth and upfront investment of VC-backed firms observed in administrative US Census data, this paper develops a firm dynamics model over the life cycle that centers on ex-ante heterogeneity in growth potential, innovation investment, and external financing. In the model, startups choose the source of financing from VC, Angel investors, and banks, where financial frictions arise from bank default costs and costs of raising equity. VC-backed firms achieve substantial growth as a result of endogenous sorting, equity-based funding, and managerial advice. The calibrated model implies that venture capitalists' advice accounts for around 22% of the growth of VC-backed firms. A counterfactual economy without VC financing would lose aggregate consumption by around 0.4%.

Keywords: Venture capital, firm dynamics, innovation, upfront investment, defaultable debt, endogenous sorting
JEL codes: D22, D25, E22, G24, G30, O32

---

[*] I am grateful to Jeremy Greenwood, Dirk Krueger, and Sergio Salgado for their guidance and helpful suggestions. I thank Harold Cole, Emin Dinlersoz, Alessandro Dovis, David Hsu, Joachim Hubmer, Mitsuru Igami, Nicolas Jacquet, Benjamin Lester, Parag Mahajan, Tatsuro Senga, Ia Vardishvili, Theodor Vladasel, and Thomas Winberry for their helpful comments. Access to the Crunchbase database is highly appreciated. All errors are my own.

[†] Boston University (Technology & Policy Research Initiative). Email: `yoando@bu.edu`. Any views expressed are those of the author and not those of the U.S. Census Bureau. The Census Bureau's Disclosure Review Board and Disclosure Avoidance Officers have reviewed this information product for unauthorized disclosure of confidential information and have approved the disclosure avoidance practices applied to this release. This research was performed at a Federal Statistical Research Data Center under FSRDC Project Number 2633. (CBDRB-FY23-P2633-R10186, CBDRB-FY23-P2633-R10510, CBDRB-FY23-P2633-R10843)

1 Introduction

Young firms play an important role in job creation, business dynamics, innovation, and economic growth.[1] In particular, a small fraction of high-growth firms make disproportionate contributions to the economy. However, startup founders with an intention to grow their business often lack the financial resources and management skills to execute their business ideas. Venture capital (VC) is a prominent example of external financing for high-growth firms and has backed more than 30% of firms that had an initial public offering (IPO) since 1980.[2] This paper examines the role that external financing, in particular VC financing, plays in capitalizing young firms and fostering high-growth firms.

Motivational facts are documented using administrative US Census data and proprietary VC datasets. In the data, VC-backed firms achieve substantial employment and payroll growth relative to firms financed by other sources and take large upfront investments. To explain the facts, this paper develops a quantitative firm dynamics model over the life cycle, in which firms are endowed with heterogeneous growth potential, choose the source of financing, and invest in innovation. In the model, high-potential firms self-select into VC financing to benefit from equity-based funding and managerial advice. The calibrated model explains the two facts reasonably well as untargeted moments. A counterfactual simulation at the micro level implies that venture capitalists' advice increases the growth of VC-backed firms by around 22% relative to Angel financing.[3] A counterfactual economy without VC financing would lose aggregate output and consumption by around 0.2% and 0.4%, respectively, in the steady state.

The empirical analysis is based on the Longitudinal Business Database at the US Census Bureau, which I merged with proprietary VC datasets by name and address matching. This facilitates a comprehensive study of VC-backed firms in the US economy. At the macro level, a growth accounting exercise shows that VC-backed firms contributed to 10.6% of employment growth and 15.8% of payroll growth for the period from 1990 to 2019. This occurred even though only around 0.2% of all firms in the economy raised

---

[1] See, for instance, Haltiwanger et al. (2016), Akcigit and Kerr (2018), and Acemoglu et al. (2022). Decker et al. (2014) document that newborn firms and high-growth firms account for around 20% and 50% of gross job creation in the US economy, respectively.

[2] Venture capital is defined as "independent, professionally managed, dedicated pools of capital that focus on equity or equity-linked investments in privately held, high growth companies" (Gompers and Lerner, 2001). See Gompers (1994), Kenney (2011), and Metrick and Yasuda (2021) for the history and characteristics of venture capital. Examples of VC-backed firms include but are not limited to Amazon, Apple, Facebook, Google, Microsoft, Moderna, Netflix, Starbucks, Tesla, and Uber (Source: SDC New Issues).

[3] Venture capitalists raise funding primarily from institutional investors, such as pension funds, in the form of limited partnerships and actively engage in monitoring and advising the firms they invest in. Angel investors are wealthy individuals who invest their own money and are generally less active in monitoring and advising firms than venture capitalists. See, for instance, Fenn et al. (1997). Angel investors do not include friends or relatives (Wetzel Jr, 1983).

VC financing.

At the micro level, around 22% of VC-backed firms born from 1980 to 2010 become high-growth firms measured in employment in contrast to less than 1% among all firms. High-growth firms, so-called Gazelles, are defined as firms whose annual average employment growth rate in the first five years is above 20% and whose employment exceeds 100 by age 20.[4] In the sample of Gazelles, the growth of VC-backed firms is exceptional. The average firm size of VC-backed Gazelles at age 15 is more than five times larger than that of non-VC-backed Gazelles.

In addition, the large upfront investments of VC-backed firms and the cost of VC financing are observed in the data. VC-backed firms typically raise funding more than 10 times their revenue at age 0, as opposed to around 3.4 times the revenue of Angel-backed firms. R&D intensity, measured by R&D expenditures over firm revenue, of VC-backed firms is around 14% higher than that of non-VC-backed firms, and it declines over age. These two facts indicate that, compared to non-VC-backed firms, VC-backed firms raise significantly larger funding relative to their revenue and invest more heavily in innovation during their initial stages. Finally, VC investors acquire around 3.3% extra equity stakes compared to Angel investors in observationally comparable financing deals, which I interpret as compensation for VCs' managerial advice.

Based on the descriptive evidence, a quantitative firm dynamics model with three key elements is developed. First, firms invest in innovation to realize their heterogeneous growth potential in contrast to a firm dynamics model with an exogenous productivity process. The optimal size of innovation investments depends on the firm's growth potential. Second, startups choose their source of financing from VC, Angel investors, and banks. Financial frictions arise from bank default costs and costs of raising equity. Because high-potential firms invest intensively in innovation and would have a high chance of defaulting on bank loans, the model exhibits endogenous sorting between firms' potential and sources of financing. Third, venture capitalists (VCs) provide managerial advice to enhance the chance of successful innovation.[5] Because VCs demand extra equity stake, startups seek VC financing only if the benefits of VCs' advice exceed the cost of giving up an additional equity stake. The value-adding effect is calibrated to match the extra equity stake VCs obtain relative to Angel investors observed in the data. To the best of my knowledge, this paper provides the first quantitative framework for explaining firm growth over the life cycle that embeds the three interconnected factors: innovation, choice of financing, and the role of VCs.

In the model, firms are born with heterogeneous growth potential under complete information and invest in innovation to move up the inherent productivity ladders. At age 0, firms earn zero revenue yet pay for labor, capital, and innovation expenditures, imply-

---

[4] Alternatively, Gazelles can be defined by firm payroll. 36% of VC-backed firms became payroll Gazelles in the data. Gazelles account for more than half of net aggregate growth in employment and payroll, as shown in Section 2. The definition of Gazelles in this paper closely follows that of Sterk et al. (2021).

[5] VCs typically serve as a board of directors and occasionally recruit executives. The model assumes that the only difference between VCs and Angel investors is the presence of managerial advice.

ing the need for external financing. Firms choose either VC, Angel, or bank financing, depending on their inherent growth potential.

In bank financing, firms take defaultable short-term debt each period. The bank loan rate is endogenously determined to satisfy the bank's zero expected profit condition based on the probability of paying back the debt and potential default costs.[6] In equity (VC or Angel) financing, firms raise funding each period in exchange for financiers' ownership stakes. The equity stake of investors is determined based on the amount of money invested and the firm's value. While there is no default with equity financing, operating a business as an equity-backed firm incurs fixed costs every period.

VC financing differs from Angel financing in that VCs provide managerial advice and improve the probability of moving up the productivity ladder, while there is a limited supply of VCs in the economy. In exchange for their advice, VCs demand an extra equity stake. The compensation for VCs is determined in equilibrium so that firms at the threshold are indifferent between VC financing and Angel financing. The 3.3% extra equity stake demanded by VCs is informative about the value-adding effect, as VC-backed firms could raise Angel financing instead if the compensation is too expensive relative to the benefit of managerial advice.

The model results in a monotone relationship between a firm's potential and the mode of financing. The majority of firms borrow from banks as their default rate is relatively low and the fixed costs of equity financing outweigh the benefits. However, bank financing is not suitable for financing high-potential firms due to their large innovation expenditures and uncertain innovation outcomes, which lead to a high likelihood of bankruptcy with bank financing and potentially high bankruptcy costs. The top 0.2% of newborn firms raise VC financing as they benefit the most from the VCs' advice due to the high return of successful innovation.[7]

The calibration targets salient features of firm dynamics. A novel part is an identification of the innovation function, in which the probability of successful innovation depends on innovation expenditures and other firm characteristics. In practice, VC-backed firms typically raise multiple rounds, and reaching a subsequent round is seen as an indication of successful business performance. The model targets the probability of raising subsequent VC rounds observed in the data, which are informative about the success rate of innovation. This approach identifies the innovation function without relying on data on R&D expenditures or patent filings, which take positive values for only a small fraction of all firms.[8] Other calibration targets include the average firm size over the life cycle,

---

[6] All defaults in bank loans correspond to Chapter 7 bankruptcy, resulting in the shutdown of the business. The default procedure incurs default costs.

[7] The model is calibrated so that 0.2% of all firms raise VC financing and 2% of all firms raise equity financing. The value-adding effect of VCs is quantified to be consistent with the 3.3% extra equity stakes of venture capitalists in equilibrium.

[8] Mezzanotti and Simcoe (2023) document that only around 6% of the overall population of US firms are active in R&D. In the PetentsView database, around 15,000–30,000 firms per year were granted at least one patent during the period from 2010 to 2020, out of around 5 million firms in the economy, implying

the coefficient on the Pareto tail in the firm-size distribution, and the firm exit rate. The default costs are set to match the loan recovery rate.

The calibrated model explains the substantial growth and upfront investment of VC-backed firms as untargeted moments. It also reasonably matches the evolution of equity stakes owned by investors and the total money raised by VC-backed and Angel-backed firms over the life cycle as untargeted moments.

The calibrated model is used to quantify the role of VCs at the micro and macro levels. At the micro level, counterfactual growth paths of VC-backed firms are simulated under alternative sources of financing. The counterfactual simulation implies that VCs' advice enhances the expected firm size of VC-backed firms at age 10 by around 22% relative to Angel financing. On the other hand, the expected firm size of VC-backed firms would decline by another 53% with bank financing relative to Angel financing, demonstrating the importance of equity financing for high-growth firms. At the macro level, the role of VCs in the economy is quantified by simulating an economy without VC financing. The counterfactual experiment predicts a decline in aggregate output and consumption in the steady state by around 0.2% and 0.4%, respectively. This implies that the impact of VCs on the economy is sizable. In addition, doubling the number of VCs in the economy would enhance aggregate output and consumption by 0.1% and 0.15%, respectively, as more firms receive VCs' advice while paying lower compensations to VCs.

Related Literature This paper contributes to four strands of the literature.

First, this paper is closely related to recent studies of venture capital in the macroeconomy with general equilibrium analysis. For instance, Jovanovic and Szentes (2013), Opp (2019), Akcigit et al. (2022), and Greenwood et al. (2022) examine the outperformance of VC-backed firms in the macroeconomy. Compared to their work, my study embeds VC financing into a firm dynamics model and explains the growth trajectories of VC-backed firms over the life cycle. In addition, I model the endogenous choice of financing from VC, Angel financing, and bank financing and analyze the benefits and costs of each source of financing.[9]

Second, the quantitative model builds on a large body of firm dynamics literature in the style of Hopenhayn (1992). The analysis echoes three themes in the literature: (i) firm dynamics over the life cycle, as in Sterk et al. (2021), (ii) financial frictions in the style of Cooley and Quadrini (2001),[10] and (iii) the productivity ladder, as in Cole et al.

---

that less than 1% of all firms generate patents each year. The innovation expenditures in the model are meant to capture broader innovation activities.

[9] In Jovanovic and Szentes (2013), entrepreneurs choose between going solo or VC financing. In Opp (2019), implementing superior ideas requires VC fund intermediation. In Akcigit et al. (2022) and Greenwood et al. (2022), the choice of VC or non-VC is exogenous.

[10] In Cooley and Quadrini (2001), firms decide production inputs one period in advance. Firms face financial frictions because of the costs of raising equity and the costs of defaulting on debt. Recent work on firm dynamics with financial frictions includes Buera and Shin (2013), Khan et al. (2014), Midrigan and Xu (2014), Ottonello and Winberry (2020), Salgado (2020), Sui (2020), Corbae and D'Erasmo (2021),

(2016). Compared to Sterk et al. (2021), the life cycle dynamics of firms with various sources of financing are tracked. Compared to Cooley and Quadrini (2001), the current analysis allows for a synergy effect between VCs and entrepreneurs. In comparison with Cole et al. (2016), the probability of moving up the ladder is endogenous and is affected by the source of financing.

Third, my work builds on the empirical literature on VC financing. Studies that estimate the treatment effect of VC include Kortum and Lerner (2000), Sørensen (2007), Chemmanur et al. (2011), Bertoni et al. (2011), and Bernstein et al. (2016). Compared to their studies, the value-adding effect of VCs relative to Angel investors is quantified using a new calibration strategy based on the extra equity stake of VC investors compared to Angel investors.[11] In addition, the growth of VC-backed firms is measured in terms of employment and payroll, which are relevant to the welfare of households.

Finally, this study contributes to the empirical literature on firm dynamics that analyzes US Census data (e.g., Haltiwanger et al., 2013; Decker et al., 2016; Pugsley and Șahin, 2019; Salgado et al., 2019; and Kondo et al., 2023). The study by Puri and Zarutskie (2012) is the first to document the growth of VC-backed firms using Census data. Compared to their study, my contribution is to document the substantial growth of VC-backed firms even among high-growth firms and the large upfront investments by venture capitalists. In addition, the evidence on (i) the contribution of VC-backed firms to aggregate growth and (ii) the high R&D intensity of VC-backed firms is new to this study.[12] This adds to the understanding of VC-backed firms in the sample of all employer businesses in the United States.

2 Descriptive Evidence

Three sets of descriptive evidence are presented. First, VC-backed firms exhibit high growth in employment and payroll in the US economy. At the macro level, VC-backed firms account for only around 0.2% of all firms yet make a sizable contribution to aggregate growth in employment and payroll. At the micro level, VC-backed firms have a significantly higher chance of becoming Gazelles than non-VC-backed firms. Conditional on becoming Gazelles, VC-backed Gazelles achieve substantial growth in comparison with non-VC-backed Gazelles. Second, VC-backed firms make large upfront investments

---

Boar et al. (2022), Kochen (2022), and Gomes and Sarkisyan (2023)

[11] Related to the extra equity stakes of VCs documented in this study, Hsu (2004) finds that entrepreneurs are more likely to accept financing offers from more reputable VCs even with a discount on the company's valuation, suggesting the extra value of reputable VCs over less-reputable VCs. Hochberg et al. (2007) show that startups invested by better-networked VCs perform better. Other related work includes Catalini et al. (2019), who estimate an upper bound to the returns to venture capital using the data on founders' choices. Gompers et al. (2020) present new survey results about how venture capitalists make decisions.

[12] Akcigit and Kerr (2018) show a negative correlation between innovation intensity, measured in patents per employee, and firm size. Ma (2022) documents the difference in R&D intensities between small-/medium and large firms.

in their early life. The R&D intensity is higher than non-VC-backed firms and declines over firm age. Third, VC investors acquire more equity stakes than Angel investors in observationally comparable financial deals. These facts motivate a firm dynamics model with VC financing and innovation. The extra equity stakes are used to quantify the role of VCs.

2.1 Data

The primary source of firm-level microdata is the Longitudinal Business Database (LBD), which is collected and organized by the US Census Bureau.[13] The LBD covers virtually the entire population of employer firms and contains information on employment, payroll, revenue, firm age, industry (NAICS code), and firm death. The sample spans from 1980 to 2019 in this study. Firm size is primarily measured in employment and payroll. Revenue data have been added to the LBD since 1997. The LBD is linked to the Census Bureau's Business Register, which contains the name and address of all establishments.

The Survey of Business Owners (SBO) in 2007 supplements the information on the source of financing. The survey asks more than one million businesses in the United States about their sources of capital, such as early-stage equity investments and business loans. The data on R&D expenditures are drawn from the Survey of Industry Research and Development (SIRD) for the period 1980–2007 and from the Business R&D and Innovation Survey (BRDIS) for the period 2008–2019. These datasets are a part of the restricted-use database of the Census Bureau and are linked to the LBD with census identifiers.

I implemented name and address matching between the Business Register and the four external VC datasets: Crunchbase, Pitchbook, Preqin, and Refinitiv's VentureXpert.[14] The VC datasets provide information on the timing and size of most VC deals in the sample and the equity share of investors for a subset of VC deals. Crunchbase and Pitchbook contain information on Angel deals, which enables comparisons between VC deals and Angel deals.[15] In addition, two datasets, the SDC New Issues and Ritter's IPO dataset, are merged to supplement the VC-backing status of firms that eventually went public.

---

[13] Appendix A.1 describes each dataset more in detail. See also Jarmin and Miranda (2002) and Chow et al. (2021) for a detailed description and the construction of the LBD.

[14] The Business Register is linked to the LBD by establishment identifiers. Crunchbase granted me access to its database for research purposes (URL: `www.crunchbase.com`). The other three VC datasets are accessed from the Wharton Research Data Services (WRDS) and the Penn Libraries.

[15] Since Angel investments are undertaken by individual investors, academic researchers generally face challenges in analyzing comprehensive data on Angel financing.

	Fraction	Employment growth	Payroll growth
VC-backed	0.2%	10.6%	15.8%
“Gazelles”	0.8%	65.2%	51.3%

Notes: Gazelles are firms with an average annual growth rate above 20% in the first five years and employment of 100 or above by age 20. The fraction of firms is computed for the period from 1980 to 2019. The contribution to aggregate growth is for 1990–2019. Data: LBD and VC datasets.

Table 1: Contribution of VC-backed Firms and Gazelles to Aggregate Growth

2.2 Growth Accounting

An accounting exercise is conducted to measure the contribution of VC-backed firms and high-growth firms to aggregate growth. Exploiting the sample of essentially all firms in the US economy, their contributions to net job creation and payroll growth are computed.[16]

First, high-growth firms, which are often called “Gazelles” in the literature, are defined based on the firm’s employment or payroll. Employment Gazelles are firms whose average annual employment growth rate is above 20% in the first five years and whose employment exceeds 100 by age 20. Payroll Gazelles are defined likewise.[17] VC-backed firms are those perceived to raise VC financing in the VC datasets and are linked to firms in the LBD. For the period from 1980 to 2019, around 0.2% of all firms are VC-backed, and around 0.8% of all firms are Gazelles, as shown in Table 1.[18]

Next, the contributions of VC-backed firms and Gazelles to aggregate growth are computed. The growth rate of all firms is computed based on the growth rates of es-

---

[16] The importance of VC-backed firms in the aggregate economy has been documented in the literature (cf. Lerner and Nanda, 2020; Greenwood et al., 2022). However, the literature typically examines publicly traded firms, which account for a relatively small share of the total employment. According to Dinlersoz et al. (2018), the domestic employment of actively-traded firms accounts for around 26% of the total private employment in the US.

[17] Payroll Gazelles are firms with an average annual payroll growth rate above 20% in the first five years and with their payroll exceeding 100 times the average real wage per worker in the economy by age 20. The definition of Gazelles closely follows Sterk et al. (2021), who define Gazelles as “startups with an ex ante projected growth rate of at least 20 percent annually, over the first five years, and an expected employment level of at least 10 workers at some point during their lifetimes.” In their definition, they find that 5.4% of all startups are Gazelles. Here, I set 100 employees instead of 10 employees to narrow the focus on high-growth firms.

[18] VC-backed firms, employment Gazelles, and payroll Gazelles account for 0.195%, 0.793%, and 0.788% of all firms, respectively. The definition of Gazelles requires observing the firm growth rate from age 0 to 5 and firm size up to 20 (unless the firm reaches employment of 100 or above before age 20). Therefore, some firms, especially those born before 1980 or after 2015, can be misclassified as non-Gazelle firms. I use the average growth rate from 1980 to 1985 to define the growth rate in the first five years for firms born before 1980.

tablishments belonging to the firms, following Davis et al. (1996) (hereafter, DHS).[19] The DHS firm growth rate is computed as the weighted sum of growth rates of establishments belonging to the firm in the next period. The advantage of this approach is that it measures the organic expansion of firms, which is not affected by ownership changes. If a firm acquires another firm, the firm size increases mechanically, but the DHS growth rate is unaffected as long as the size of the establishments belonging to the firm does not change. Based on the DHS firm growth rates, the growth rate (g_{st}) of a group of firms is computed as a weighted sum of the firm growth rates within the group s. Then, it is straightforward to calculate the contribution of group s to aggregate growth in a particular year t:

$$\tilde{g}_{st} := \left(\frac{X_{st}}{X_t} \right) g_{st},$$

where $\left(\frac{X_{st}}{X_t} \right)$ represents the employment or payroll share of group s in the economy. The contribution of groups of firms (i.e., VC-backed firms or Gazelles) over a long period of time is derived by taking the ratio between the geometric mean of the gross rates of contribution $1 + \tilde{g}_{st}$ and the mean of aggregate growth rates $1 + g_t$.[20]

The results of the growth accounting are summarized in Table 1. Despite their small fraction in the economy, VC-backed firms account for 10.6% of employment growth and 15.8% of payroll growth from 1990 to 2019. Gazelles account for 65.2% of employment growth and 51.3% of payroll growth. Their contributions are disproportionally large. In terms of payroll growth, VC-backed and Gazelle firms make approximately 80 and 65 times more contributions to aggregate growth, respectively, than typical firms in the economy. Note that VC-backed firms contribute more to payroll growth than to employment growth, implying that VC-backed firms create more high-paying jobs.

	Any firm	Employment Gazelles	Payroll Gazelles
All firms	18,150,000	156,000 (0.86%)	160,000 (0.88%)
VC-backed firms	27,500	6,000 (21.8%)	9,900 (36%)

Notes: Gazelles are firms with an average annual growth rate above 20% in the first five years and employment of 100 or above by age 20. Data: LBD and VC datasets.

Table 2: New firms from 1980 to 2010 in the LBD

2.3 High Growth of VC-backed Firms

The firm's growth at the micro level is analyzed in the sample of firms born from 1980 to 2010. Approximately 18.2 million firms were created during this period (Table 2). By linking with the VC datasets, 27,500 firms are identified as VC-backed.[21] Thus, around 0.15% of newborn firms are backed by VCs during this period.[22] Employment Gazelles and payroll Gazelles account for 0.86 and 0.88% of newborn firms, respectively. Notably, 21.8% and 36% of VC-backed firms become employment Gazelles and payroll Gazelles, respectively. The odds are around 25 to 40 times higher than typical firms in the economy.

---

[19] See also Haltiwanger et al. (2013). An establishment-level DHS growth rate g_{it} is defined as:

$$g_{it} = \frac{E_{i,t} - E_{i,t-1}}{X_{it}}, \text{ where } X_{it} = \frac{1}{2}\left(E_{i,t-1} + E_{i,t}\right),$$

where $E_{i,t}$ denotes employment or payroll in establishment i at time t.

The firm growth rate g_{ft} is a weighted sum of the growth rates of establishments that belong to the same firm:

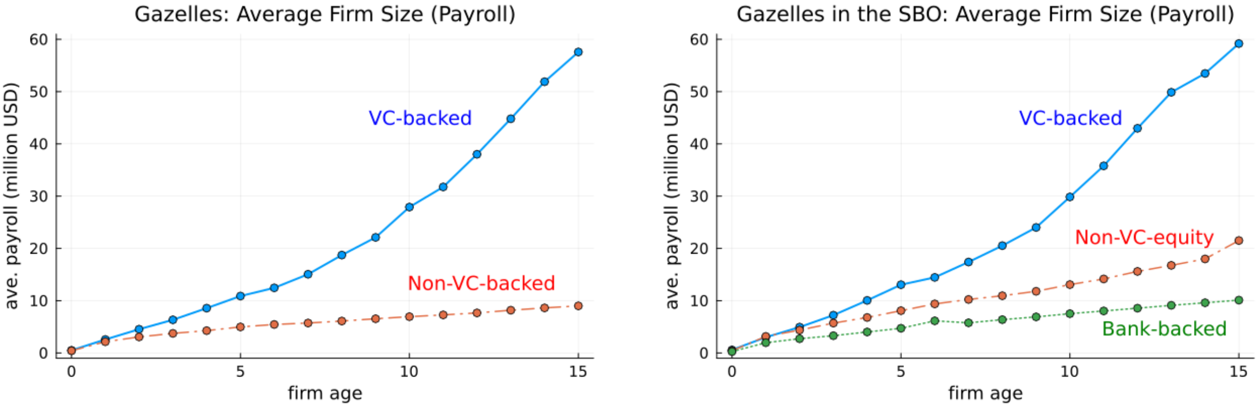
$$g_{ft} = \sum_{i \in f} \left(\frac{X_{it}}{X_{ft}}\right) g_{it}, \text{where } X_{ft} = \sum_{i \in f} X_{it}.$$

The size of firm f is denoted by X_{ft}, and the size of establishment i that belongs to firm f is denoted by X_{it}. The growth rate of a group of firms g_{st} or the aggregate growth rate g_t is a weighted sum of the growth rates of firms belonging to the same group or all firms in the economy.

[20] For example, if the long-run aggregate growth rate is 2% and the long-run contribution by group s is 0.2%, the group's share in aggregate growth is 10%.

[21] The number of firms is properly rounded following the disclosure avoidance rules of the US Census Bureau.

[22] Here, 0.15% is lower than 0.2% of VC-backed firms among all firms between 1980 and 2019 documented in Table 1. This is because the fraction of VC-backed firms is higher after 2010. In this subsection, firms born from 1980 to 2010 are analyzed so that the growth rate from age 0 to 5 and firm size for at least 10 years are observed for all firms. Compared to the literature, Puri and Zarutskie (2012) report 17,763 VC-backed firms and 16,286,524 non-VC-backed firms born from 1981 to 2005. This implies that about 0.1% of newborn firms are VC-financed. I obtained slightly different statistics for three reasons: (a) more firms received VC financing in more recent years, (b) I merge the four VC datasets instead of two datasets in their study, and (c) I partially address the issue of spurious new firms, as described in Appendix A.1.



VC vs. non-VC-backed　　VC, Non-VC-Equity, and Bank

Notes: The sample of Gazelle firms born from 1980 to 2010. The sample in the right panel is a subset of firms that reported the source of capital in the Survey of Business Owners (2007). Payroll is denominated in 2009 US dollars. Data: LBD (1980-2019), SBO(2007), and VC datasets.

Figure 1: Growth Trajectory of VC-backed and non-VC-backed Gazelles

Thus, VC-backed firms are much more likely to become Gazelles than non-VC-backed firms.

Next, in the sample of Gazelles, the growth trajectories of VC-backed and non-VC-backed Gazelles are compared. The left panel in Figure 1 shows the average firm size of VC-backed and non-VC-backed Gazelles from age 0 to 15, measured in payroll. VC-backed firms sustain their high growth, whereas the growth of non-VC-backed Gazelles is moderate.[23] The average firm size of VC-backed Gazelles at age 15 is around \$60 million, more than five times that of non-VC-backed Gazelles.

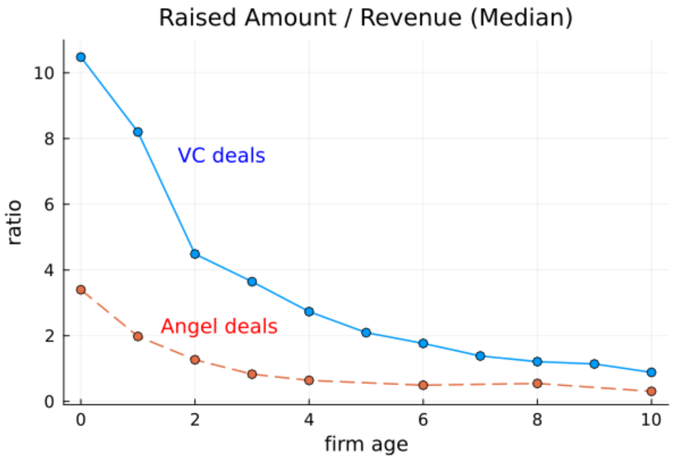
The right panel compares the growth trajectories of VC-financed, non-VC-equity-financed, and bank-financed Gazelles. The information on the source of capital is drawn from the SBO (2007). This figure visually shows the stark difference in growth trajectories depending on the source of financing.

Note that the descriptive evidence about the high growth of VC-backed firms is documented here without showing the causal relationship. The quantitative model will be developed to explain the high growth of VC-backed firms as a result of endogenous sorting, equity-based funding, and managerial advice. The model is also used to predict the growth path of firms with an alternative source of financing.[24]

---

[23] Regression Table 10 in Appendix A.2 shows that the DHS firm growth rate is higher for VC-backed Gazelles than non-VC-backed Gazelles, conditional on the initial year and NAICS code. See Figures 13 and 15 in Appendix A.2 for the robustness of this pattern in employment and revenue. A comparison of exit rates is shown in Table 11. See Puri and Zarutskie (2012) for a comparison unconditional on the high-growth status. The GDP deflator is used to convert payroll into constant 2009 US dollars. The average employment of VC-backed Gazelles at age 15 is above 1,000, as can be seen in Figure 13.

[24] One approach to disentangle the value-adding effect from the selection effect would be to control for as many observable characteristics as possible. However, relatively little information is available about



Notes: The raised amount is drawn from the VC datasets. Revenue is from the LBD (1997-2019).

Figure 2: Upfront Investment of VC-backed and Angel-backed Firms

2.4 Upfront Investment

In the previous subsection, a stark difference in firm growth is documented depending on the source of financing. Upfront investments of VC-backed firms are examined in this subsection, which helps understand the underlying mechanism.

The ratio of the size of funding raised in VC or Angel financing to the firm's revenue is documented at each age from 0 to 10. The median in the sample of VC-backed and Angel-backed firms is depicted in Figure 2.[25] This shows that VC-backed firms typically raise more than 10 times their revenue at age 0. The funding size consistently exceeds the revenue until age 9. Angel-backed firms also raise funding larger than their revenue at age 0, but the ratio decreases to less than 1 by age 3. VC-backed firms make a particularly large upfront investment in their early life compared to Angel-backed firms. Note that such large upfront investments cannot be explained with a standard firm dynamics model that does not consider innovation activities.

In addition to the evidence of upfront investments based on the size of equity investments, the upfront investments by VC-backed firms using the data on R&D expenditures is documented, which is a commonly used measure of innovation inputs. The R&D data are drawn from the SIRD/BRDIS datasets at the US Census. The two datasets cover a large sample of firms investing in R&D and contain young firms that are not publicly traded.[26] The R&D intensity is defined as R&D expenditures divided by revenue. The

---

newborn firms aside from firm age and industry. In addition, the qualities of newborn firms are intrinsically unobservable. The Deaton-Paxson regression (Figure 15) in Appendix A.2 shows that the difference in growth trajectories is not mainly driven by year or industry effects.

[25] To avoid disclosing confidential information, the pseudo median, which is the mean of a subset of observations along the median, is reported following the guidelines of the Census Bureau. In the sample of Angel-backed firms, observations at ages 4-5, 6-7, 8-9, and 10-11 are combined to secure a reasonable sample size.

[26] See Ottonello and Winberry (2023), for example, for the analysis of R&D expenditures made by pub-

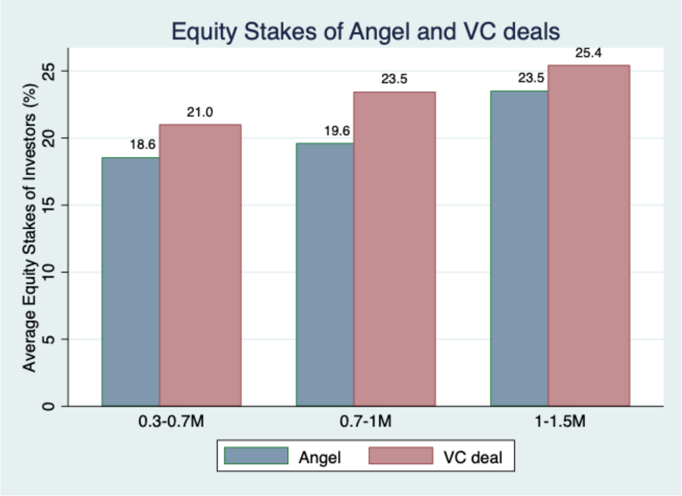
	Dependent: R&D/revenue (%)	
	(1) all firms	(2) payroll Gazelles
VC-backed dummy	14.5 ***	14.2***
	(0.743)	(0.882)
firm age	–0.0629***	–0.0788***
	(0.0137)	(0.0204)
VC × age	–0.347***	–0.306***
	(0.0391)	(0.0459)
log(emp)	–1.46***	–1.24***
	(0.0684)	(0.121)
Year	Yes	Yes
NAICS (4 digits)	Yes	Yes
N	241,000	80,000
R^2	0.329	0.436

Standard errors in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Notes: Firms with positive R&D are in the sample. R&D/revenue are winsorized at 5%. The sample in column (2) is restricted to payroll Gazelles. The data on R&D expenditures are taken from the SIRD (1980-2007) and BRDIS (2008-2019) datasets. Observations are weighted by the sampling weights. The sample period is from 1980 to 2019.

Table 3: R&D/revenue of VC-backed and non-VC-backed firms



Notes: The first financing deals (either VC or Angel deals) for each company are included in the sample. Sample period: 1990-2021. Sample size: 4135. Data source: Pitchbook.

Figure 3: Equity Stake in VC deals and Angel deals

regression results comparing the R&D intensity of VC-backed and non-VC-backed firms are reported in Table 3.[27] It shows that VC-backed firms typically have higher R&D intensity by around 14 percentage points relative to non-VC-backed firms and that the intensity declines over firm age. A similar pattern is observed in the sample of Gazelles in column (2). Those results show that VC-backed firms intensively spend on innovation, particularly in their early life. To my knowledge, the evidence of upfront investment of VC-backed firms, measured by the raised amount over revenue or R&D expenditures over revenue, is new to this study.

2.5 Equity Share in VC Financing

This subsection documents the difference in equity stakes obtained by VC investors and Angel investors. This evidence implies that VCs' advice is valuable from the startup's perspective based on the revealed preference argument. The empirical moment will be a crucial input to the model, as it helps to disentangle the perceived value-adding effect of VCs. Figure 3 provides a first glance at the average equity stakes in VC and Angel financing. In the sample of first deals, VC investors acquire more equity stakes than Angel investors conditional on deal size categorized into three bins. The difference ranges from 1.9% to 3.9%, depending on the size bin.

Then, regression analysis is employed to control for the characteristics of firms observed in the LBD, such as log(employment), firm age, and industry (Table 4). Column (1) shows that the size of VC deals is typically larger than that of Angel deals. Column (2) shows that VC investors acquire around 3.3% extra equity stakes conditional on ob-

---

licly traded firms.

[27] Firms with positive R&D expenditures are analyzed in the regression.

	Dependent variable:		
	(1) log(deal size)	(2) equity share (%)	(3) deal size/revenue
VC deal dummy	1.448***	**3.285*****	7.232***
log(employment)	0.502***	–3.873***	
log(deal size)		5.731***	
firm age			–0.994***
Year	Yes	Yes	Yes
Age Dummy	Yes	Yes	No
Deal No	Yes	Yes	Yes
Initial Year	Yes	Yes	Yes
NAICS (4 digits)	Yes	Yes	Yes
N	37500	22500	37500

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Notes: (deal size/revenue) is winsorized at 5%. Data: LBD (1997-2019) and the VC datasets.

Table 4: Comparison between VC deals and Angel deals

servable firm characteristics.[28] This regression coefficient will be used in the calibration of the model. Column (3) confirms the evidence on upfront investment documented in Figure 2: VC-backed firms raise larger financial capital relative to their revenue than Angel-backed firms, and the ratio declines over firm age.

Finally, Table 5 reports the firm age at the first VC financing. Almost half of the VC-backed firms obtain the first VC deal by age 0, and around three-fourths of VC-backed

---

[28] One may be concerned about firms' qualities unobservable to researchers. If VC-backed firms have higher unobserved qualities, the extra equity stakes in the regression can be underestimated, as higher-quality firms would give up smaller equity stakes. If 3.3% extra equity stakes is underestimated, then the value-adding effect of VCs would be even greater than that predicted by the calibrated model.

	Fraction (%)	Cumulative Fraction (%)
Age $\leq$ 0	46.77	46.77
1$\leq$ Age $\leq$ 2	27.45	74.22
3$\leq$ Age $\leq$ 5	14.05	88.27
Age $>$ 5	11.73	100

Data: LBD (1980-2019) and the VC datasets.

Table 5: Age at the First VC deal

firms raise VC funding by age 2.[29] This motivates the construction of the model in which all firms make a choice of financing at birth.

3 The Model

A firm dynamics model with ex-ante heterogeneity, innovation, and external financing is developed to explain the high growth of VC-backed firms and to quantify the role of VCs. New firms are born with heterogeneous growth potential and choose the source of external financing at birth.[30] The choice of financing and innovation expenditure forms the course of a firm's growth trajectory. Firm dynamics over the life cycle are studied in general equilibrium.

3.1 Production

Each firm has inherent growth potential, which can be thought of as the quality of business ideas. An ambitious idea has the potential to yield high growth, but it does not materialize with certainty. The potential is represented by a sequence of productivities $\{\hat{\theta}_1, \cdots, \hat{\theta}_N\}$, which is called a productivity ladder. Firms invest in innovation to move up the productivity ladder. The probability of moving up the ladder (successful innovation) depends on innovation expenditures, VCs' advice, and other firm characteristics. The firm achieves its full potential when it reaches the last step of the productivity ladder. The firm also faces an exogenous productivity shock independent of its growth potential. This results in nontrivial dynamics of firm size and survival status.

Firms are risk-neutral and maximize the discounted sum of future expected profits. Firms produce goods with a decreasing return-to-scale technology by hiring workers and renting capital. The production function is:

$$y_t = \underbrace{(\hat{\theta}_{n(t)} + \epsilon_t)^{1-\eta}}_{\text{productivity}} \left(k_t^{\alpha} \ell_t^{1-\alpha}\right)^{\eta}, \tag{1}$$

where k_t and ℓ_t denote capital and labor, respectively. Capital and labor are static inputs and purchased at prices R and w, respectively. The costs are paid one period in advance before generating revenue. A span-of-control parameter is given by $\eta < 1$.

A firm's productivity is a sum of two elements, $\hat{\theta}_{n(t)}$ and ϵ_t. The first element $\hat{\theta}_{n(t)}$ is determined by the firm's potential and innovation outcome. All firms start with the first

---

[29] Firm age is 0 in the year when the oldest establishment in the firm reports positive payrolls for the first time.

[30] The model assumes that information is complete and abstracts from informational issues, such as incomplete information and asymmetric information. This approach has limitations since venture capitalists may be better at screening startups. Note that Sørensen (2007) estimates a two-sided matching model of venture capital and distinguishes between VCs' influence and sorting under the assumption of complete information. Ewens et al. (2022) develop and estimate a search-and-matching model in which entrepreneurs and venture capitalists are informed of each other's type upon encounter.

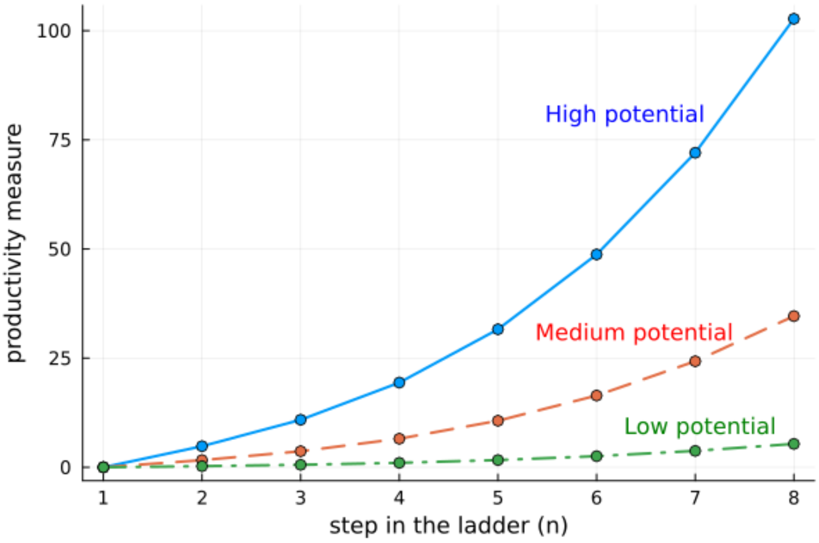


Figure 4: Examples of Productivity Ladder

step of their productivity ladder, $n = 1$. If their innovation is successful with probability ς, they move up to the next step. Otherwise, the firms stay at the same step of the ladder in the next period. This is formalized in the following equation:

$$n(t) = \begin{cases} n(t-1)+1 & \text{with prob. } \varsigma_t \\ n(t-1) & \text{with prob. } 1-\varsigma_t \end{cases}, \tag{2}$$

where $n(t) \in \{1, ..., N\}$ denotes a position in the ladder at time t. The probability of moving up the ladder ς_t is endogenously determined by innovation expenditures. The second element ϵ_t follows an exogenous AR(1) process:

$$\log(\epsilon_t) = \rho \log(\epsilon_{t-1}) + \sigma_\nu \nu_t \text{ with } \nu_t \sim \text{Normal}(0, 1). \tag{3}$$

Figure 4 depicts examples of the productivity ladder. Each firm is born with a sequence of productivities $\{\hat{\theta}_n\}_{n=1,\cdots,N}$. Different productivity ladders capture heterogeneity in business ideas. Some founders start their businesses with an ambitious idea and are inspired to sell their products to a large number of customers in the future. Most founders pursue a less ambitious business idea and have little intention of expanding their business. Nonetheless, the idea is merely a blueprint at birth, and all firms start from an initial step of the ladder $n = 1$ with $\hat{\theta}_1 = 0$. They enhance productivity by succeeding in innovation and climbing up the ladder.

The probability of moving up the ladder $\varsigma_{t+1} \in (0, 1)$ is determined in the following logistic function:

$$\underbrace{\varsigma_{t+1}}_{\text{success rate}} = \frac{\exp(x_t)}{1+\exp(x_t)}, \tag{4}$$

$$\begin{aligned} \underbrace{x_t}_{\text{factor of success}} &= \gamma_0 + \underbrace{\gamma_1}_{(+)} \underbrace{\log(I_t)}_{\text{innovation expnd}} + \underbrace{\gamma_2}_{(-)} \underbrace{\log(1+\hat{\theta}_N)}_{\text{firm's potential}} \\ &\quad + \underbrace{\gamma_3}_{(-)} \mathbf{1}\{\text{fail}_t\} + \underbrace{\gamma_4}_{(+)} \mathbf{1}\{\text{VC}\} + \gamma_5 \underbrace{\log(n_t)}_{\text{step in the ladder}}, \end{aligned} \tag{5}$$

where the success rate ς_{t+1} increases with the factor of success x_t. The factor of success (x_t) is assumed to depend on six terms in equation (5) to match empirical moments discussed in Section 4.1.[31] The second term means that the success probability is increasing in innovation expenditure I_t. The third to sixth terms in equation (5) are other firm characteristics relevant to the success rate. A negative coefficient γ_2 means that more ambitious ideas with higher potential $\hat{\theta}_N$ are harder to materialize. A negative γ_3 implies that if a firm failed in innovation last period, the probability of success also declines in the coming period. In practice, startups that are not growing will have a hard time keeping their momentum.[32] A coefficient γ_4 measures the value-adding effect of VCs. As VCs provide managerial advice, the firm has a higher chance of successful innovation. Finally, a coefficient γ_5 captures the relation between the success probability of innovation and the position on the ladder. All six coefficients are estimated in calibration.

Firms obtain revenue from production only in the next period, as in Cooley and Quadrini (2001), and need to raise external funding if their cash flow is short of the expenditure. In practice, a critical job of startup founders is to avoid running out of cash, which is a serious concern, especially for firms that choose a path of high growth and a rapid scale-up of the business.[33] The setup for the timing of expenditure and revenue in the model captures this practical aspect.

Figure 5 shows the timeline. Firms are born with growth potential $\hat{\theta} := \{\hat{\theta}_n\}$, initial exogenous productivity ϵ_0, and initial assets o_0. Given their potential, firms decide the

---

[31] If the factor of success (x_t) depends only on the second term (i.e., $\gamma_0 = \gamma_2 = \gamma_3 = \gamma_4 = \gamma_5 = 0$), innovation expenditures (I_t) can be expressed as a function of innovation success rate:

$$I_t = \left(\frac{\varsigma_{t+1}}{1-\varsigma_{t+1}}\right)^{1/\gamma_1}. \tag{6}$$

This functional form is similar to Greenwood et al. (2022), where the innovation expenditures can be expressed as $I_t = \text{constant} \times \frac{\varsigma_{t+1}^2}{1-\varsigma_{t+1}}$. Both functional forms have the advantage of giving an interior solution for innovation expenditures.

[32] For instance, Michael Seibel, managing director of Y Combinator, says about startups: “The second that you launch, every week that you’re not growing, you’re dying.” Source: “7 Rules Every Tech Startup Must Follow,” *Fortune*, November 3,2016. In Cole et al. (2016), where they introduce the productivity ladder, if a firm stalls at the current step of the ladder, it remains at the current level forever after, implying $\gamma_3 = -\infty$ in this setting. I relax this assumption and set γ_3 in calibration.

[33] See, for instance, Feld and Ramsinghani (2013) and Feld and Mendelson (2019), for issues that startup founders face in starting their business and managing financing.

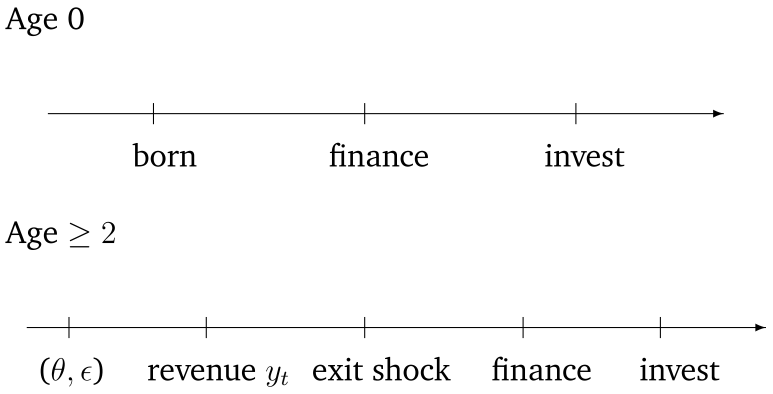


Figure 5: Timeline

source of funding: VC financing, Angel financing, or bank financing. Then, they raise money and pay for labor, capital, and innovation expenditures. In the first period, at age 0, firms do not collect revenue from their production.

At the beginning of period t, productivity $(\hat{\theta}_{n(t)}, \epsilon_t)$ is realized. Given labor and capital inputs (l_t, k_t) chosen in the previous period, the firm earns revenue according to the production function (1). After the realization of revenue, firms may be hit by an exogenous exit shock with probability φ. In this case, firms cannot continue the business and exit the market. Otherwise, firms may raise additional external funding, pay for labor, capital, and innovation expenditures, and produce goods in the next period. If the business is not profitable enough to operate, the firm may endogenously exit the market. This may happen as firms need to pay for operational fixed costs c_f every period. One period is set to be two years in the model. The details of the firm's decisions are discussed in the following subsections.

3.2 Equity (VC or Angel) Financing

Firms collect revenue only in the next period after paying for labor input, capital input, and innovation expenditures and thus may need to rely on external financing. Firms raise money from VC investors, Angel investors, or banks. The choice of the source of financing depends on the firm's initial characteristics.[34]

In equity financing, firms take financial investment from investors in exchange for an equity stake in the company. Firms do not have an obligation to pay back the investors until they earn profits. Since investors own a part of the firm, they will receive dividends once the business profits materialize in the future. The downside of equity financing in the model is that firms pay a fixed cost $c_{e,0}$ every period to operate as equity-backed

---

[34] In the model, firms make the choice at birth and cannot raise money from other sources. Practically, firms may raise money from several sources at the same time or over time. I assume only one source of financing to maintain computational tractability. Hellmann et al. (2021) find that VC financing and Angel financing are dynamic substitutes and that firms are less likely to switch between the two.

firms.[35] Additionally, raising one unit of money costs $c_{e,1}$ as flotation costs. Those two costs are the distortions of equity financing.

Venture capitalists provide managerial advice. The advice is beneficial for enhancing the probability of moving up the productivity ladder. In exchange, firms pay management fees $\bar{\pi}_{\text{vc}}$ every period to the venture capitalists.[36] The supply of VCs in the market is fixed, and the payments to the VCs are determined in equilibrium. The compensation $\bar{\pi}_{\text{vc}}$ is reflected in the VCs' extra equity stakes. In contrast, Angel investors operate in a competitive market and make zero expected profits. Except for the advice and management fees to VCs, VC financing and Angel financing have the same structure.

Equity investors take the equity share (s_t) of the firm based on the value of the financial investment relative to the firm's value.[37] Firms may raise money in several stages, and in the second or later round of equity financing, the investor's share in earlier rounds will be diluted. Taking this into account, the total equity share of investors s_{t+1} evolves according to the following formula:[38]

$$\underbrace{s_{t+1}}_{\text{total stakes}} = \underbrace{s_t}_{\text{previous stakes}} + (1 - s_t) \underbrace{\frac{(1 + c_{e,1})(-d_t)}{V_t(n_t, \epsilon_t, o_t, \text{fail}_t; \hat{\theta}) + (1 + c_{e,1})(-d_t)}}_{\text{newly acquired stakes}}, \tag{7}$$

where $(-d_t)$ denotes the amount of funding raised in this period, and $V_t(n_t, \epsilon_t, o_t, \text{fail}_t; \hat{\theta})$ denotes the firm's value before raising money.

A value function of equity-backed firms, $V_t(n_t, \epsilon_t, o_t, \text{fail}_t; \hat{\theta})$, is now formulated. The four state variables are the position of the ladder (n_t), the current exogenous productivity shock (ϵ_t), cash flow at the beginning of the period (o_t), which is equal to the firm's revenue in equity financing except for the initial period, and the most recent innovation outcome ($\text{fail}_t \in \{0, 1\}$). Given the state variables, firms decide whether to continue their business and, if so, choose labor and capital inputs and innovation expenditures. The revenue minus the sum of all expenditures is distributed as dividends.

Firms exit the market if the continuation value is smaller than their cash flow. Hence, the firm's value is the maximum of the continuation value and its cash flow:

$$V_t^j(n_t, \epsilon_t, o_t, \text{fail}_t; \hat{\theta}) = \max\{\underbrace{V_t^{j,C}(n_t, \epsilon_t, o_t, \text{fail}_t; \hat{\theta})}_{\text{continue}}, \underbrace{o_t}_{\text{cash flow}}\} \text{ for } j \in \{\text{VC}, \text{Angel}\}. \tag{8}$$

---

[35] Equity-backed firms typically register as incorporated businesses, establish a board of directors, hire lawyers and accountants, and disclose financial information to their investors. The cost of raising equity includes these costs.

[36] In reality, firms and VCs form a sophisticated contract in which payment to VCs is contingent on the firm's performance. The setup here simply captures that VCs are compensated for their managerial advice.

[37] Intuitively, the investor's share s_t is equal to the value of money invested divided by the firm's post-money valuation. The two are not exactly the same in the model because of the flotation costs and compensation for VCs.

[38] See Appendix B.1 for the derivations. Note that equity investments increase the firm's value by the amount of money invested in the firm. The value of past investors' equity stake does not change because of the money injection. Hence, the equity share of the past equity investors will decrease to offset the increase in the firm's value.

If the firm continues the business operation, the continuation value is as follows:

$$V_t^{j,C}(n_t,\epsilon_t,o_t,\text{fail}_t;\hat{\theta}) = \max_{\ell_{t+1},k_{t+1},I_t,d_t} \underbrace{d_t}_{\text{dividend(+) or deal size(–)}} - \underbrace{\mathbf{1}\{(-d_t)>0\}(-d_t c_{e,1})}_{\text{cost of raising equity}} \tag{9}$$

$$+\frac{1}{1+r}\left\{(1-\varphi)\underbrace{\mathbb{E}_{n_{t+1},\epsilon_{t+1}}\left[V_{t+1}^{j}(n_{t+1},\epsilon_{t+1},o_{t+1},\text{fail}_{t+1};\hat{\theta})|\varsigma_{t+1}\right]}_{\text{no exit shock}} + \varphi\underbrace{\mathbb{E}_{n_{t+1},\epsilon_{t+1}}\left[o_{t+1}|\varsigma_{t+1}\right]}_{\text{exit shock}}\right\}$$

$$\text{s.t. } d_t + \underbrace{Rk_{t+1}+w\ell_{t+1}+I_t+c_f+c_{e,0}+\mathbf{1}\{\text{j=VC}\}\bar{\pi}_{\text{vc}}}_{\text{expenditures}} = \underbrace{o_t}_{\text{cash flow}} \tag{10}$$

$$o_t = (\hat{\theta}_{n_t}+\epsilon_t)^{1-\eta}(k_t^{\alpha}\ell_t^{1-\alpha})^{\eta}, \tag{11}$$

where d_t is the value of the dividend if it is positive, and $(-d_t)$ is the amount of money raised if d_t is negative.[39] If the firm is hit by an exogenous exit shock in the next period with probability φ, the firm must exit the market with its cash flow o_{t+1}. Otherwise, the firm value is given by equation (8). In equation (10), R is the rental rate of capital, w is the wage of workers, I_t is the innovation expenditure, c_f is the fixed operational costs, and $c_{e,0}$ is the additional fixed cost for equity-backed firms. In VC financing, firms pay management fees $\bar{\pi}_{\text{vc}}$ to VCs every period. Innovation expenditures I_t determine the probability of stepping up the ladder ς_{t+1} in the next period, according to the innovation functions (4) and (5). If the sum of all expenditures is larger than revenue, firms rely on external financing.

Management fees to venture capitalists $\bar{\pi}_{\text{vc}}$ is determined in equilibrium. Since the supply of VCs is limited, only a fixed fraction of all startups can be funded by VCs.[40] Hence, startups need to compensate for VCs' advice. While newborn firms are heterogeneous in their potential, VCs are homogeneous. Thus, all VCs obtain the same compensation. In equilibrium, all VC-backed firms compensate $\bar{\pi}_{\text{vc}}$ for VCs' advice so that the threshold firm is indifferent between VC financing and Angel financing.[41] In calibration, $\bar{\pi}_{\text{vc}}$ is determined such that 0.2% of newborn firms choose VC financing.

The equity share of investors follows equation (7). The maximization problem above does not involve the equity share of investors. Appendix B.1 shows that this formulation is equivalent to maximizing the founder's value, which is given by the equity share of the founder multiplied by the firm's post-money valuation. The compensation for VCs is reflected in the extra equity stakes, as $\bar{\pi}_{\text{vc}}$ increases the amount of money raised $(-d_t)$

---

[39] In the model, firms never distribute dividends and raise money from investors at the same time, as external financing is more costly than internal financing. In equity financing, firms do not make a savings decision in the model. This is mainly for computational tractability, although the savings decision is less relevant for young firms, especially for firms scaling up their business.

[40] Jovanovic and Szentes (2013) also assume a limited supply of VCs.

[41] If startups pay less than π_{vc}, VCs will find other startups that are willing to accept higher compensation. If VCs demand more than π_{vc}, startups will find other VCs willing to accept slightly lower compensation. The stable match can be microfounded as an equilibrium under a proper market mechanism, such as a deferred acceptance algorithm. I thank Nawaaz Khalfan for the helpful discussion.

in equation (10), resulting in higher equity stakes in equation (7).[42]

3.3 Bank Financing

Bank-backed firms raise defaultable short-term debt instead of raising equity. The advantage of bank financing in the model is that it does not entail fixed costs ($c_{e,0}$) or flotation costs ($(-d)c_{e,1}$), as in equity financing. However, if a firm cannot pay back debt, it goes bankrupt. This procedure incurs default costs, and banks seize only a part of the firm's assets. Banks take into account the possibility of a firm's bankruptcy and the associated default costs and charge a bank loan rate such that the expected return of the loan is equal to the risk-free rate.

A firm raises debt b_{t+1} with price $q_t(b_{t+1}, k_{t+1}, I_t; n_t, \epsilon_t, o_t, \text{fail}_t, \hat{\theta})$ at time t and pays back b_{t+1} in the next period. The price of debt satisfies $q_t < 1$ and is a function of control variables (b_{t+1}: size of debt, k_{t+1}: capital input, I_t: innovation expenditures) and the firm's state variables. The bank loan rate is given by the inverse of the price $1/q_t$ and is equal to the risk-free rate, $1 + r$, if the default probability in the next period is zero.

In bank financing, the firm's cash flow is revenue minus the debt repayment:

$$o_t = y_t - b_t. \tag{12}$$

If the cash flow is negative, the firm either defaults on debt, in which case the firm loses all revenue and shuts down the business, or borrows additional money and rolls over the debt. The decision to default, denoted by $\Delta_{t+1} = 1$, is given by:

$$\Delta_{t+1} = \mathbf{1}\left\{y_{t+1} < b_{t+1} \text{ and } (V_{t+1} < 0 \text{ or exit shock})\right\}. \tag{13}$$

The firm chooses to default on the debt if the outstanding debt exceeds revenue and the firm does not operate in the next period, either because of the firm's negative present value or because of an exogenous exit shock.

Banks operate in a competitive market and charge a bank loan rate that satisfies the following zero-expected-profit condition:

$$q_t(\cdot) = \frac{1}{1+r}\mathbb{E}_t\left[\underbrace{\mathbf{1}\left\{\Delta_{t+1} = 0\right\}}_{\text{no default}} + \underbrace{\mathbf{1}\left\{\Delta_{t+1} = 1\right\}}_{\text{default}}\left(\frac{y_{t+1}}{b_{t+1}} - c_d\right)\right]. \tag{14}$$

If a firm does not default on debt ($\Delta_{t+1} = 0$), a bank receives full repayment on debt. In case of default ($\Delta_{t+1} = 1$), a return on debt for the bank is given by $\left(\frac{y_{t+1}}{b_{t+1}} - c_d\right)$. This is because the firm transfers all revenue (y_{t+1}) to the bank upon default, but the default incurs costs that are proportional to the size of the debt.

A value function of a bank-backed firm is formulated. At the beginning of each period, firms choose either to continue their business, pay back their debt and exit, or default on the loan and exit. Thus, the firm's value is the maximum among the continuation value,

---

[42] Intuitively, as startups are short of cash flow to pay for VCs, they give up equity stakes to VCs instead.

the value of cash flow, and the value of default, which is zero from the business owner's perspective.

$$V_t^{\text{Bank}}(n_t, \epsilon_t, o_t, \text{fail}_t; \hat{\theta}) = \max\{\underbrace{V_t^{\text{Bank},C}(n_t, \epsilon_t, o_t, \text{fail}_t; \hat{\theta})}_{\text{continue}}, \underbrace{o_t}_{\text{exit}}, \underbrace{0}_{\text{default}}\}. \tag{15}$$

The continuation value is given by the following:

$$\begin{aligned}
V_t^{\text{Bank},C}(n_t, \epsilon_t, o_t, \text{fail}_t; \hat{\theta}) &= \max_{d_t \geq 0, \ell_{t+1}, k_{t+1}, I_t, b_{t+1}} d_t \qquad (16)\\
&+ \frac{1}{1+r}\left\{(1-\varphi)\underbrace{\mathbb{E}_{n_{t+1},\epsilon_{t+1}}\left[V_{t+1}(n_{t+1}, \epsilon_{t+1}, o_{t+1}, \text{fail}_{t+1}; \hat{\theta})|\varsigma_{t+1}\right]}_{\text{no exit shock}} + \varphi \underbrace{\mathbb{E}_{n_{t+1},\epsilon_{t+1}}\left[\max\{0, o_{t+1}\}|\varsigma_{t+1}\right]}_{\text{exogenous exit}}\right\}\\
\text{s.t. } d_t + \underbrace{Rk_{t+1} + w\ell_{t+1} + I_t + c_f}_{\text{expenditures}} &= \underbrace{o_t}_{\text{cash flow}} + \underbrace{q_t(b_{t+1}, k_{t+1}, I_t)b_{t+1}}_{\text{new debt}} \qquad (17)\\
o_t &= (\hat{\theta}_{n_t} + \epsilon_t)^{1-\eta}(k_t^{\alpha}\ell_t^{1-\alpha})^{\eta} - b_t. \qquad (18)
\end{aligned}$$

with the price of debt q_t and the default decision Δ_{t+1} given by equations (13) and (14), respectively. The probability of moving up the ladder ς_{t+1} is determined by innovation functions (4) and (5). If the cash flow is larger than total expenditures, the firm distributes the business surplus as a dividend. Otherwise, firms borrow money from a bank or cut expenditures to meet their budget constraints. The optimality conditions and numerical algorithms to compute the value function iteration are described in Appendix B.2.4.

3.4 Choice of Financing

Firms with heterogeneous initial characteristics choose the source of financing that maximizes the firm's present discounted value:[43]

$$V_0(n_0, \epsilon_0, o_{\text{initial}}, \hat{\theta}) = \max\left\{V_0^{\text{VC}}(n_0, \epsilon_0, o_{\text{initial}}, \hat{\theta}), V_0^{\text{Angel}}(n_0, \epsilon_0, o_{\text{initial}}, \hat{\theta}), V_0^{\text{Bank}}(n_0, \epsilon_0, o_{\text{initial}}, \hat{\theta})\right\}. \tag{19}$$

Figure 6 shows an example of the firm's value in the model depending on growth potential. In the figure, high-potential firms attain the highest value if they choose VC financing. On the other hand, low-potential firms obtain the highest value with bank financing. Intuitively, high-potential firms benefit the most from VCs' advice since the gain from successful innovation is exceptionally high.[44] Low-potential firms do not heavily invest in innovation, so the probability of default and the associated default costs in bank

---

[43] Since the firm's value already takes into account the payment to VCs, it can be interpreted as the firm value from the owner's perspective.

[44] In the model, productivity increase due to successful innovation is larger for high-potential firms than low-potential firms. Fox et al. (2012) estimate a two-sided matching game between entrepreneurs and VCs and show complementarities.

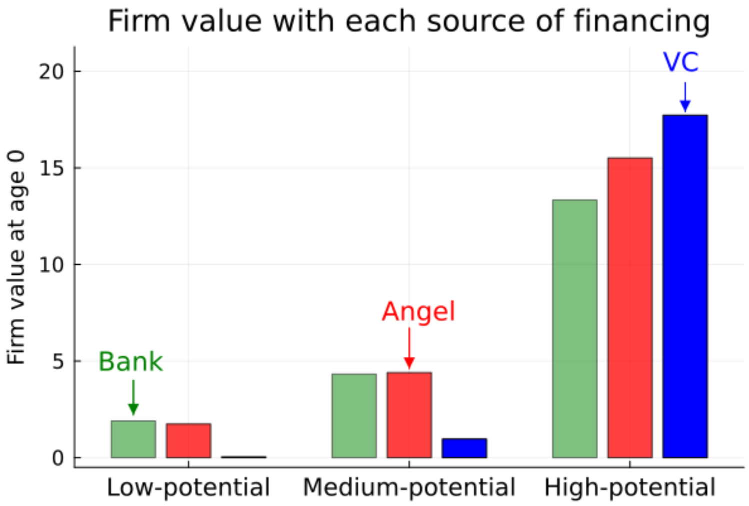


Figure 6: The Choice of Financing at Age 0

financing are low.[45] Medium-potential firms choose Angel financing as equity financing is more suitable than bank financing to finance innovation, but the compensation for VCs is too expensive. This results in a monotone relationship between a firm's potential and the mode of financing.

Firms are born with three initial characteristics $(\hat{\theta}, \epsilon_0, o_0)$. The growth potential $\hat{\theta} = \{\hat{\theta}_n\}_{n=1,\cdots,N}$ is summarized by a single number $\hat{\theta}_N$, which is the productivity at the last step of the ladder. Productivity levels in other rungs of the ladder $\{\hat{\theta}_n\}_{n=1,\cdots,N}$ are determined by the following parametric function of $\hat{\theta}_N$ and n:

$$\hat{\theta}_n = \hat{\theta}_N \left[(1-\xi) \left(\frac{n-1}{N-1} \right) + \xi \left(\frac{n-1}{N-1} \right)^3 \right]. \tag{20}$$

This functional form satisfies $\hat{\theta}_1 = 0$ and $\hat{\theta}_N = \hat{\theta}_N$. If $\xi = 0$, the firm's productivity increases linearly along the ladder. If $\xi > 0$, the ladder has a convex shape. The curvature of the ladder ξ is calibrated in Section 4.1 to match average firm growth in the data.[46] The growth potential $\hat{\theta}_N$ follows a log-normal distribution:[47]

$$\hat{\theta}_N \sim \text{LogNormal}(\mu_\theta, \sigma_\theta). \tag{21}$$

---

[45] Figure 16 in Appendix B.3.1 shows the relation between a firm's potential and default rate.

[46] The cubic term is introduced to allow for a steep, non-linear shape of the ladder. The quadratic term is dropped to reduce the number of parameters in calibration and to avoid a colinearity issue.

[47] To numerically solve the model, I draw nine values of $\hat{\theta}_N$'s from the log-normal distribution at the percentiles: [0.2, 0.5, 0.7, 0.85, 0.935, 0.9725, 0.9775, 0.989, 0.999]. The probability measures assigned to each percentile are $\frac{1}{100}\times$[40, 20, 20, 10, 7, 0.5, 0.5, 1.8, 0.2], respectively.

The initial exogenous productivity ϵ_0 also follows a log-normal distribution:[48]

$$\epsilon_0 \sim \text{LogNormal}(\mu_{\epsilon 0}, \sigma_{\epsilon 0}). \tag{22}$$

The initial assets o_{initial} are assumed to be the same for all firms.[49]

Free Entry Condition: Entrants are ex-ante identical. Their growth potential $\hat{\theta}$ and exogenous productivity ϵ_0 are observed upon entry. The free entry condition ensures that the ex-ante value of the entry is equal to the cost of entry denoted by f^e:

$$\text{Value of Entry} := \sum_{\epsilon_0, \hat{\theta}} \underbrace{V_0(n_0, \epsilon_0, o_{\text{initial}}; \hat{\theta})}_{\text{firm value given an initial state}} \Phi(\epsilon_0, \hat{\theta}) = \underbrace{f^e}_{\text{cost of entry}}, \tag{23}$$

where the growth potential $\hat{\theta}$ and initial exogenous productivity ϵ_0 follow distributions (21) and (22). $\Phi(\epsilon_0, \hat{\theta})$ denotes the probability mass of firms with initial states $(\epsilon_0, \hat{\theta})$. All firms start from an initial step of the ladder ($n_0 = 1$) and with initial assets o_{initial}.[50]

3.5 Aggregation

This section describes a representative household and capital renters. The economy is closed with market clearing conditions.

Representative Household: The representative household supplies a fixed amount of labor $\bar{L}$ every period. All profits from heterogeneous firms and venture capitalists are returned to the household, denoted by Π_t.[51] The infinitely lived household maximizes the lifetime utility subject to a budget constraint:

$$\max_{c_t, a_{t+1}} \sum_{t=0}^{\infty} \beta^t u(c_t) \tag{24}$$

$$\text{s.t. } c_t + a_{t+1} = w_t \bar{L} + (1+r)a_t + \Pi_t, \tag{25}$$

$$\text{where } \Pi_t = \sum_i \left[d_{it} + \mathbf{1}\{\text{VC}_i\}\bar{\pi}_{\text{vc}}\right]. \tag{26}$$

Each firm is indexed by i. The household's saving a_t returns a risk-free rate, $1 + r$. Flow utility is assumed to be CRRA (constant relative risk aversion). The household makes

---

[48] The AR(1) process of ϵ in equation (3) is approximated by the Tauchen method. Appendix B.3.2 describes how grid points are defined. The number of grid points for ϵ is set to five. Since production does not take place in the initial period, ϵ_0 is relevant only as indicative information for the exogenous productivity ϵ_1 in the next period.

[49] In practice, initial assets may be correlated with the growth potential $\hat{\theta}_N$ and initial productivity ϵ. However, initial assets at the firm level are typically not observed in the data.

[50] Newborn firms have not failed in innovation, and their failure status starts with $\text{fail}_0 = 0$

[51] Households pay for initial assets of newborn firms and receive cash flows of firms hit by an exogenous exit shock.

the consumption and saving decision, following the Euler equation:

$$u'(c_t) = \beta(1+r)u'(c_{t+1}). \tag{27}$$

In the steady state, the Euler equation implies:

$$1 + r = \frac{1}{\beta}. \tag{28}$$

Hence, the steady-state risk-free rate is given by the inverse of the discount factor β.

Capital Renters: Capital renters collect savings from the representative household and rent capital to firms with a rental rate R. As capital depreciates with rate δ, the rental rate of capital R is given by:[52]

$$R = r + \delta. \tag{29}$$

The capital rental market is competitive, and capital renters make zero profits.

Market Clearing Conditions: The labor market, capital market, and final goods market clear in equilibrium. First, the sum of the labor demand of heterogeneous firms is equal to the exogenous labor supply of the household:

$$\sum_i l_{it} = \bar{L}. \tag{30}$$

Second, aggregate capital rented by capital owners is equal to the household's savings:

$$K_{t+1} = a_{t+1}. \tag{31}$$

Finally, the total output produced plus undepreciated capital is equal to the sum of all expenditures, household consumption, and capital in the next period:

$$\underbrace{\sum_i y_{it}}_{\text{output}(=:Y_t)} + (1-\delta)K_t = \underbrace{\sum_i \left[I_{it} + c_f + f^e\mathbf{1}\{\text{age}_i = 0\} + \mathbf{1}\{\text{equity}_i\}(c_{e,0} + c_{e,1}(-d_{it})) + \Delta_{it}c_d b_{it}\right] + C_t + K_{t+1}}_{\text{all expenditures}(=:\text{EXPND}_t)}.$$

In the steady state, capital in the economy is invariant over time. Therefore, the household's consumption in the steady state is given by:

$$C = Y - \text{EXPND} - \delta K. \tag{32}$$

The distribution of heterogeneous firms with state variables $(n_i, \epsilon_i, o_i, \text{fail}_t; \hat{\theta}_i)$ is stationary in the steady state.[53] In the quantitative analysis, wages are set to match the average firm size over the life cycle. In the steady state, the rental price of capital R is determined by the discount rate of the household and the depreciation rate, as in equations (28) and (29). The price of final goods is normalized to one. The measure of entrants in the steady state is given by M^e.

---

[52] The maximization problem can be formulated as:

$$\max_{K_t} \, (R + 1 - \delta)K_t - (1+r)K_t.$$

[53] Appendix B.3.3 describes how the stationary distribution is computed numerically.

4 Quantitative Analysis

The model is calibrated to capture the salient features of firm dynamics and explains the substantial growth and upfront investments of VC-backed firms reasonably well as untargeted moments. The calibrated model is used to quantify the importance of VC financing at the micro and macro levels.

4.1 Calibration

The model features 23 parameters listed in Table 6. Nine parameters are fixed exogenously. The remaining 14 parameters are internally calibrated to minimize the distance between model-implied moments and targeted data moments. The targeted moments are shown in Figure 7, Figure 8, and Table 7.

Exogenous Parameters: The model period is two years, so the risk-free rate is set to $r = 0.08$. Capital depreciates at rate $\delta = 0.2$ biannually. The coefficient on capital in the production function (1) is $\alpha = 0.33$. The span-of-control parameter is $\eta = 0.85$, as is standard in the literature.[54] An exogenous biannual exit rate is set to $\varphi = 0.08$.[55] The persistence parameter in the exogenous productivity process (3) is taken to be $\rho = 0.8$.[56] The standard deviation of exogenous productivity shocks σ_ν is normalized to one. The unit cost of raising capital is set equal to $c_{e,1} = 0.04$, within a range of values reported in the literature.[57] Finally, startups launch their businesses with initial assets o_{initial}, which are chosen to be equal to the wage of one worker.[58]

Fitted Parameters: The rest of the parameters are chosen to match empirical moments. The first set of parameters $(\gamma_0, \gamma_1, \gamma_2, \gamma_3, \gamma_4, \gamma_5)$ in the innovation function (5) are particularly unique to this study. Note that estimation of innovation functions is generally challenging as researchers do not directly observe innovation inputs and innovation outputs.[59] I estimate the coefficients using the characteristics of VC deals and Angel deals

---

[54] See Midrigan and Xu (2014), for example.

[55] Sterk et al. (2021) estimate an annual exogenous exit rate of around 4.1% in their calibration.

[56] Foster et al. (2008) report persistence of various TFP measures and demand shocks ranging from 0.76 to 0.97 at an annual rate. Instead of exogenously setting this parameter, one could internally calibrate ρ to be consistent with the autocorrelation of firm size in the data at the expense of computational burden. The calibrated model predicts the autocorrelation of log(employment) of around 0.8, within the range of biannual autocorrelation of around 0.75 to 0.9 at various ages, reported by Sterk et al. (2021).

[57] Gomes (2001) estimates the unit flotation costs of 0.028. Hennessy and Whited (2007) estimate the linear cost of external equity of around 0.091.

[58] Robb and Robinson (2014) report that the average owner equity of startups is around \$33,000 in 2004. The average wage in 2004 was around \$36,000, according to the Social Security Administration (URL: `https://www.ssa.gov/oact/cola/awidevelop.html`).

[59] A typical approach would be to measure innovation inputs with R&D expenditure and innovation outputs with patent filing. However, most firms in the economy do not invest in R&D and do not produce patents. Hence, this approach would significantly narrow down the scope of innovation. R&D is defined as

Parameters	Values	Description	Identification
Innovation Function			
γ_0	5.61	intercept	Pr(2nd VC round)
γ_1	0.46	coef on $\log(I)$	equity share of VC (age 0)
γ_2	-0.96	coef on $\log(1 + \hat{\theta}_N)$	equity share of Angel (age 0)
γ_3	-4.47	coef on failure	average firm size over life cycle
γ_4	0.12	coef on VC	VCs' extra equity share (3%)
γ_5	0.08	coef on $\log(n)$	Pr(5th VC round)
Productivity Distribution			
μ_θ	-4.09	mean of $\log(\hat{\theta}_N)$	fraction of Gazelles
σ_θ	3.81	std of $\log(\hat{\theta}_N)$	Pareto tail
μ_{ϵ_0}	1.68	mean of $\log(\epsilon_0)$	average firm size over life cycle
σ_{ϵ_0}	4.01	std of $\log(\epsilon_0)$	average firm size over life cycle
ρ	0.8	persistence	exogenous
σ_ν	1	std of AR(1) shocks	normalization
Financial Frictions			
$c_{e,0}$	0.29	fixed cost of equity	top 2% choose equity
$c_{e,1}$	0.04	unit cost of equity	exogenous
c_d	0.41	default costs	loan recovery rate
o_0	0.41	initial assets	wage of one worker
Other Parameters			
r	0.08	risk-free rate	standard value (2yr)
δ	0.2	capital depreciation	standard value (2yr)
α	0.33	capital share	standard value
η	0.85	span of control	standard value
φ	0.08	exit shock	exogenous (2yr)
ξ	0.69	convex ladder	average firm size over life cycle
c_f	0.36	operating costs	exit rate

Table 6: Parameters in the Model

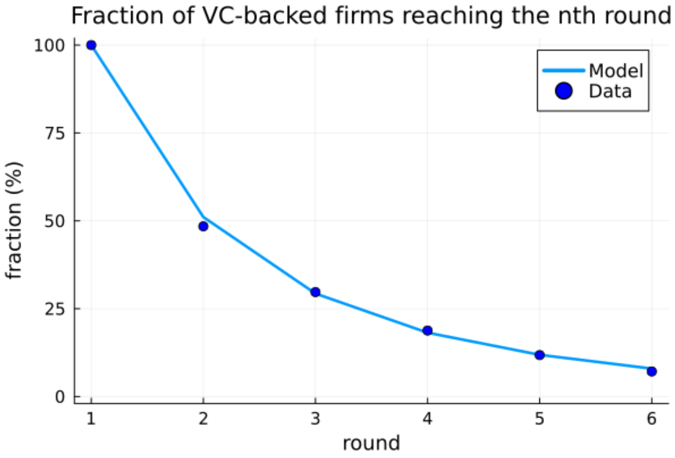


Figure 7: Probability of Subsequent VC Deals

observed in the data.

In practice, VC-backed firms raise money with multiple rounds, allowing venture capitalists to assess the firm's performance before injecting additional capital. This implies that raising additional VC rounds is an indication of satisfactory business performance.[60] Hence, I set parameters so that the model-implied probabilities of raising the second and fifth rounds hit the empirical counterparts (48.4% and 11.8%, respectively). Figure 7 shows that the model and the data match well in this dimension. This identifies the coefficients γ_0 (intercept) and γ_5 (the coefficient on $\log(n_t)$) in the innovation function, which governs the innovation success rates at the initial and upper steps of the ladder conditional on innovation expenditures.

In addition, $\gamma_3 < 0$ measures the negative impact of innovation failure on future innovation. If the magnitude of γ_3 is large enough, VC-backed firms will not raise the next VC round in case of innovation failure. This parameter also affects the average growth of all firms, which helps pin down the value.

The elasticity of the probability of successful innovation to innovation expenditures (γ_1; the coefficient on $\log(I_t)$) is pinned down based on the equity stakes of VC investors at firm age 0 and the probability of raising the second VC round. Intuitively, VC funding, equivalent to approximately 33% of firm value, is informative about innovation expenditures relative to firm values. The idea is that since VC-backed firms raise money to invest primarily in innovation and the investor's equity share is approximately equal to the money injected divided by the firm value, the equity share contains information about

---

"creative and systematic work undertaken in order to increase the stock of knowledge and to devise new applications of available knowledge," according to the BRDIS survey form. It does not include activities related to the firm's business expansion, such as market research, management studies, and the creation of new software based on known methods and applications. My study examines firm growth not necessarily based on R&D activity or patent generation.

[60] Other empirical studies, including Ewens et al. (2018), also use the next VC round as a proxy for successful business outcomes.

the size of innovation expenditures relative to firm value.[61]

The parameter γ_2 (the coefficient on $\log(1+\hat{\theta}_N)$) determines the relation between the firm's potential and innovation success rate. This parameter is pinned down by targeting the average equity stake of Angel investors in the first deal (around 23%). Intuitively, this moment is informative about the intensity of innovation activities of Angel-backed firms.[62]

The value-adding effect of VCs, denoted by γ_4, is identified by the extra equity stake of venture capitalists documented in the empirical section. The idea is to measure the perceived value-adding effect of VCs from the startup founder's perspective. The parameter γ_4 is jointly determined with the VCs' management fees $\bar{\pi}_{\text{vc}}$. Firms at the 99.8th percentile (top 0.2%) are at the threshold and indifferent between VC financing and Angel financing. While VC financing boosts the probability of successful innovation, VCs require around 3.3% of extra equity stake. The parameter γ_4 and VCs' management fees $\bar{\pi}_{\text{vc}}$ are set so that the benefit and cost balance out for firms at the top 0.2% threshold.

The second set of parameters $(\mu_\theta, \sigma_\theta, \mu_{\epsilon_0}, \sigma_{\epsilon_0})$ govern the productivity distribution. The distribution of growth potential $\hat{\theta}_N$, LogNormal$(\mu_\theta, \sigma_\theta)$, is set to match the fraction of Gazelles among all startups, 0.87%, taken from Table 2, and the Pareto tail of the firm-size distribution, 1.06, among firms with more than 100 employees, as reported by Luttmer (2010). The distribution of initial productivity ϵ_0, LogNormal$(\mu_{\epsilon_0}, \sigma_{\epsilon_0})$, regulates the average firm size in early ages.

The curvature of the productivity ladders (ξ) and wage (w) also govern the evolution of average firm size. Higher wage w lowers the average firm size, while the curvature of the ladder ξ forms the trajectory of average firm size growth. The average size of all firms from age 0 to 14 is targeted, as depicted in Figure 8. The model fits the average

---

[61] In the model, since firms also spend the raised money for labor and capital inputs and equity share reflects the VCs' profits, the relation between the equity share and γ_1 is not immediately seen. However, it is instructive to show the relation between the two under simplifying assumptions.

The FOC with respect to innovation expenditure I gives:

$$\frac{I}{\mathbb{E}\left[V_{t+1}(\text{success}) - V_{t+1}(\text{fail})\right]} = \frac{\gamma_1 \varsigma^*(1-\varsigma^*)}{1+r}, \tag{33}$$

where the firm value is given by:

$$V_t = \frac{1}{1+r}\left\{\varsigma V_{t+1}(\text{success}) + (1-\varsigma)V_{t+1}(\text{fail})\right\}. \tag{34}$$

If $V_{t+1}(\text{fail}) \approx 0$, the two equations imply:

$$\frac{I}{V_t} \approx \gamma_1(1-\varsigma^*). \tag{35}$$

Hence, if all raised money is used for innovation expenditure and if VCs make zero profits (i.e., $\frac{I}{V_t} \approx$ equity share), γ_1 can be derived in closed form, using the equity share of VC investors.

[62] Based on the approximated relation (35), one can infer that a lower equity share is related to a higher innovation success rate ς.

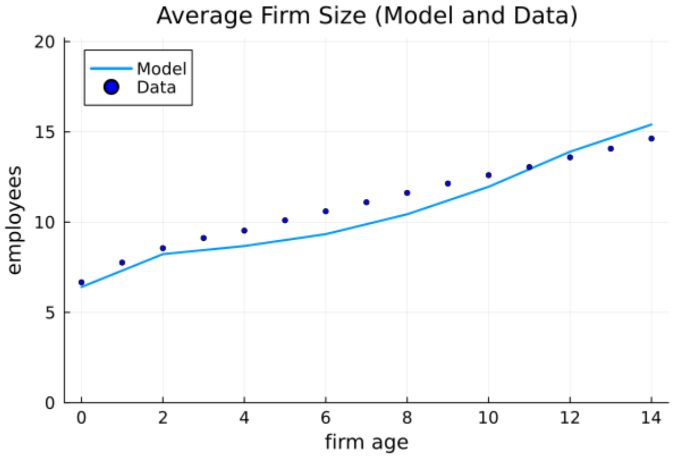


Figure 8: Average Firm Size from Age 0 to 14

firm size over the life cycle reasonably well.[63] The coefficient c_f is chosen to match the exit rate, as higher operating costs c_f induce more firms to exit.[64]

Finally, the two parameters $(c_{e,0}, c_d)$ related to financial frictions are set so that 2% of all firms choose equity financing (VC or Angel) and that banks recover 54% of defaulting loans in case of bankruptcy, a value taken from Ottonello and Winberry (2020) and Gomes and Sarkisyan (2023).[65] Parameter $c_{e,0} = 0.29$ implies that equity-backed firms pay fixed costs each period that are less than hiring one additional worker. Parameter $c_d = 0.41$ implies that defaulting on bank loans incurs costs equivalent to 41% of outstanding debt.

Overall, 19 moments are targeted to set 14 parameter values and two endogenous prices (w and $\bar{\pi}_{\text{vc}}$) in the model.[66] The calibrated parameters minimize the sum of

---

[63] The average employment at each age is taken from Sterk et al. (2021)'s replication file. I target the firm size at age (0,2,4,6,8,10,12,14), which gives eight moments in calibration.

[64] It turns out that the high exit rate observed in the data, around 9.4% among firms at age 0-20 taken from Business Dynamics Statistics, is difficult to explain with the model. Since the model period is two years, targeting an annual exit rate of 9.4% implies around 18.8% of firms exit every period, which requires large volatility in exogenous productivity shocks and high operating costs. However, this would give rise to higher growth of surviving firms and would not be in line with the average firm size in the data. One can see that Sterk et al. (2021) also struggle to explain the average firm size and exit rate jointly. A potential solution may be to introduce size-dependent operating cost shocks, as in Kochen (2022), but it is beyond the scope of this study. I prioritize explaining the average firm size, as is central to this study, and put a lower weight on the exit rate in the calibration.

[65] Sohl (2012) reports that the Angel market invested in over 10 times more deals than the VC market in the 2000s. According to the Kauffman Survey, a panel dataset of around 5,000 startups collected from 2004 to 2011, the fraction of firms that raise Angel financing at age 0 is around five times that of VC financing. In the Survey of Business Owners (2007), the fraction of non-VC-equity-backed firms is around four times that of VC-backed firms. Since we do not know the precise fraction of Angel-backed firms in the economy, I assume that 2% of all newborn firms in the model raise Angel or VC financing.

[66] I take two moments in Figure 7, eight moments in Figure 8, and nine moments in Table 7. The prices in the steady state are $w = 0.42$ and $\bar{\pi}_{\text{vc}} = 7.2$. In the baseline economy, the measure of entrants M^e is normalized to one. The calibrated model implies that the value of entry is 2.5.

Description	Target	Model
Fraction of “Gazelles”	0.87%	0.92%
Pareto tail	1.06	1.10
Average equity share at age 0 (VC)	33.0%	34.2%
Average equity share at age 0 (Angel)	23.2%	22.6%
Exit rate (annual) at age 0-20	9.4%	8.0%
Loan recovery rate	54%	54%
Fraction of VC-backed firms	0.2%	0.2%
Fraction of equity-backed firms	2%	2%
Extra equity share of VC investors	3.3%	3.3%

Notes: Gazelles are firms with an average annual growth rate above 20% in the first five years and employment of 100 or above by age 20.

Table 7: Remaining Target Moments

squared distances between targets and model-implied moments.[67]

4.2 Model Fit

The calibrated model explains the substantial growth and upfront investments of VC-backed firms documented in Section 2. The model's prediction on the evolution of equity shares and the average total money raised in VC financing and Angel financing are compared with the data. All those moments are untargeted and match their empirical counterparts reasonably well.

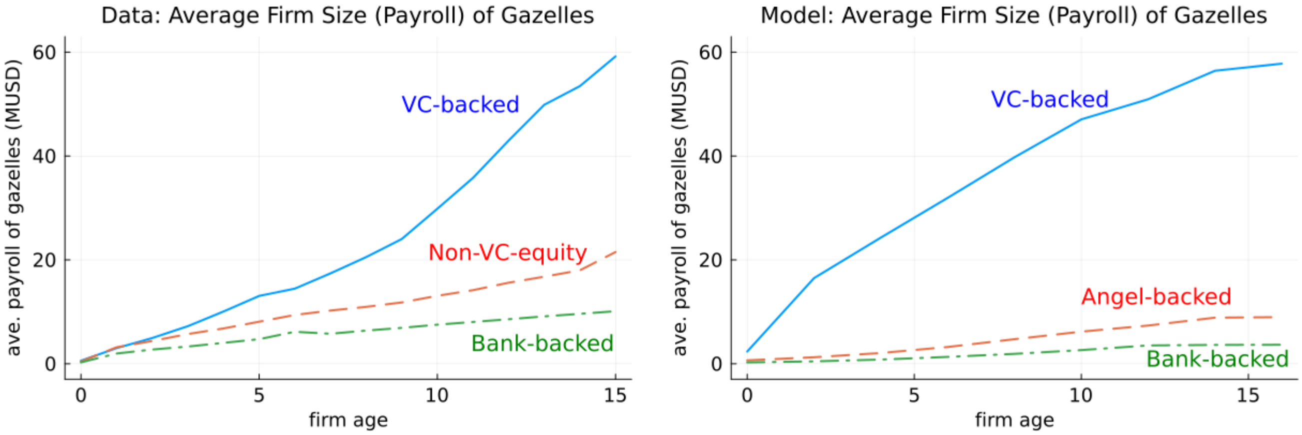
Two Motivational Facts: The first striking feature of VC-backed firms observed in the data is their substantial growth compared to non-VC-backed firms in the sample of Gazelles. To assess the model's prediction, the sample of Gazelles is taken in the model's simulation following the same definition as in the data, and the average firm size is computed at each age, separately for VC-backed, Angel-backed, and bank-backed firms. Figure 9 describes the average growth trajectories over the life cycle in the data (left panel) and in the model (right panel). It shows that the model is able to explain the substantial growth of VC-backed firms relative to non-VC-backed firms. The model is consistent with the qualitative differences among the three groups and quantitatively fits the growth of VC-backed Gazelles.[68]

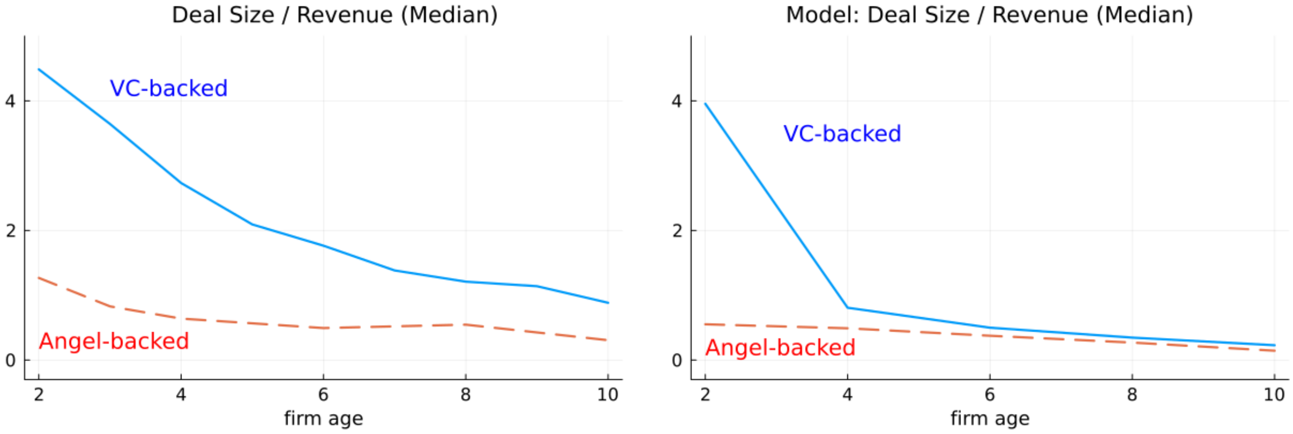
---

[67] The Nelder-Mead method is used in calibration. The eight moments about the average firm size are weighted by 5/8. In addition, the moment about the exit rate is weighted by 1/5, and the moment on the Pareto tail is weighted by 2, given the importance of the moments in the quantitative model. The other moments have the same weight as one.

[68] In calibration, I target the average firm size of all firms and the Pareto tail of the stationary firm-size distribution, neither of which directly controls the growth trajectory of Gazelles in the three groups. In

Figure 9: Growth Trajectory of Gazelle Firms (Data vs. Model)





Notes: The figures on the left show the motivational facts documented in Section 2. The figures on the right show the model's predictions.

Figure 10: Upfront Investment of High-Potential Firms (Data vs. Model)

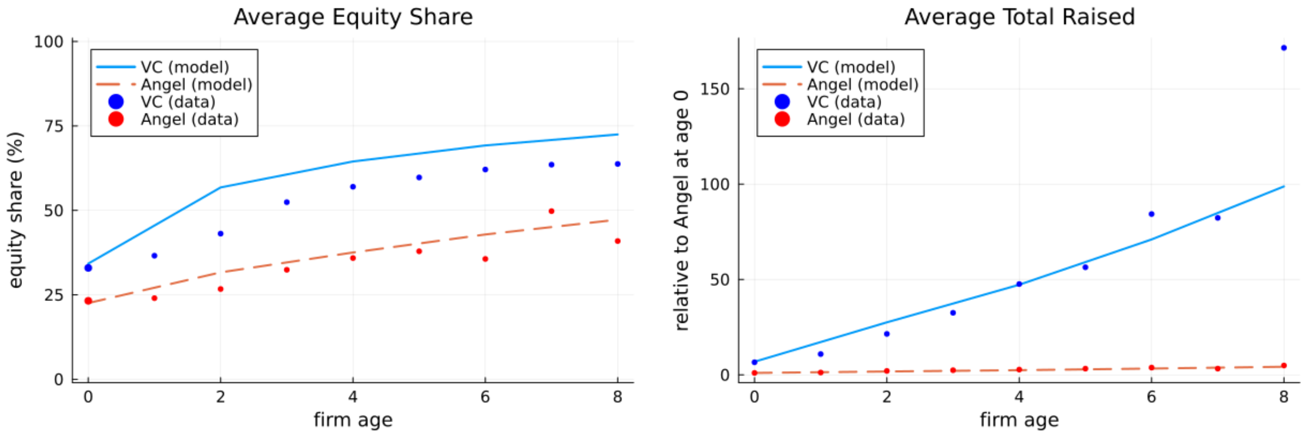
The second motivational fact is the upfront investment of equity-backed firms. To assess the model's prediction, the simulated sample of VC-backed and Angel-backed firms is taken, and the ratios of the amount of money raised to the firm's revenue are computed. The median values over the firm's life cycle are depicted in Figure 10. As in the data, VC-backed firms make large upfront investments, several times larger than their revenue, in their early life.[69]

Other Untargeted Moments: The two results above support the claim that the model provides an explanation for the substantial growth of VC-backed firms and is consistent

---

Figure 9, firm size is measured in payroll in the data. In the model, I first compute the average firm size in terms of employment. Then, I multiply the values with $40,000 to make a direct comparison with the data.

[69] Since firms earn zero revenue at age 0 in the model, the raised amount over revenue at age 0 is not depicted in the right panel.


Average Equity Share

Average Total Raised

Notes: The average equity share at age 0 is targeted in calibration, while other moments are untargeted. Equity share represents the total equity share of investors. The total money raised is shown relative to the average money raised at age 0 in Angel financing. The data are drawn from Pitchbook.

Figure 11: Average Equity Shares and Total Money Raised (Data vs. Model)

with the salient feature of VC-backed firms. On the other hand, as the model is developed based on the two motivational facts, it is ideal to test the model with other dimensions of the data.

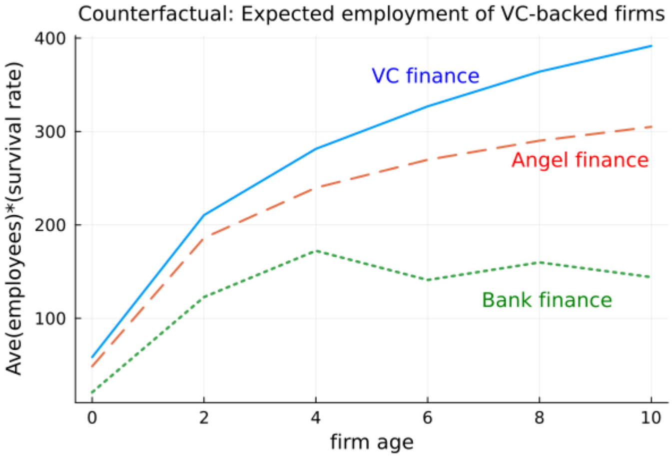
Figure 11 compares the model's prediction with the data on the evolution of the average equity share of investors and the total money raised for VC-backed and Angel-backed firms over their life cycle.[70] According to the data, VC-backed firms typically transfer more than half of their ownership stake to their investing partners after several rounds are raised. Angel-backed firms also offer additional ownership stakes over the course of financing but less so than VC-backed firms. These features are well-captured in the model (left panel). In the right panel, the total money raised from age 0 to 8 is reported. The variables are normalized by the funding size of Angel-backed firms at age 0. In the data, the amount of money raised by VC-backed firms is significantly larger than Angel-backed firms and increases substantially over the life cycle.[71] This aspect is also reasonably well-captured in the model. Finally, the annualized default rate among all firms in the model is around 0.5% and is close to a value reported in the literature.[72]

---

[70] The data are drawn from Pitchbook, and firms are observed only in the year the firm obtained new external funding. To construct the model's counterpart, I take an average of cumulative equity shares of investors and the total money raised in the sample of firms with positive external funding. In the data, I drop firms that take the first financing deal after age 2. In the model, all firms raise their first financing deal at age 0. Only a small fraction of firms above age 8 are observed in the data.

[71] In the data, Angel-backed firms raise an average of \$1.2 million at age 0 and \$4 million by age 8. VC-backed firms initially raise \$5.6 million at age 0 and \$103 million by age 8 on average.

[72] Gomes and Sarkisyan (2023) report a corporate default rate of 0.4% from Moody's. According to the Administrative Office of the US Courts, there were 24,735 business bankruptcy filings in 2015, which is around 0.4% of all firms. Source: `https://www.uscourts.gov/`.



Notes: I simulate the expected firm size (=survival rate × average employment of surviving firms) of VC-backed firms (solid blue line) in the model. Then, I simulate the firm dynamics of the same firms with alternative sources of financing (Angel or bank). The expected firm size in each case is plotted in the figure.

Figure 12: Model's Prediction on the Growth Trajectory of VC-backed Firms with Alternative Sources of Financing

4.3 Counterfactual Experiments

Using the calibrated model, the role of venture capitalists is quantified at the micro and macro levels.

In the counterfactual experiment at the micro level, the hypothetical growth path of VC-backed firms is simulated while the wage is kept unchanged. In a counterfactual economy that differs from the baseline economy in available sources of financing, the wage is determined in equilibrium to satisfy the free entry condition (23). In the steady state, the measure of entrants M^e is constant over time and is determined by the labor market clearing condition (30) written as:

$$M^e \sum_{h=0}^{\infty} \sum_{\epsilon_0,\hat{\theta}} \ell_h(n_h, \epsilon_h, o_h; \hat{\theta})\ \Phi(\epsilon_0, \hat{\theta}) = \underbrace{\bar{L}}_{\text{exogenous labor supply}} . \tag{36}$$

The total labor demand is the sum of labor demand by each cohort $h = 0, 1, \cdots$. Firms in cohort $h = 0$ are age 0 in the economy. Firms with $h = 1$ were born one period before and are at age 2, since one period in the model is two years.

The role of VCs: The value-adding effect of VCs' managerial advice is identified in the calibration based on the extra equity share of VC deals. The contribution of VCs' advice at the micro level is quantified by predicting the growth trajectories of VC-backed firms if they instead raise funding from Angel investors or banks. Figure 12 shows the

	Without VCs	Without any equity	More VCs
consumption (EV)	–0.43%	–0.76%	0.15%
output	–0.16%	–0.50%	0.10%
wage	–0.18%	–0.43%	0.04%
emp share of gazells	–5.71%	–6.32%	2.14%

Notes: I compare counterfactual economies in which some sources of financing are not available. Without the VC sector, all firms raise either Angel or bank financing. Without equity financing, all firms rely on bank financing. The percentage change in consumption is identical to the equivalent variation (EV), as the household's utility depends only on consumption. The equilibrium wage clears the free entry condition.

Table 8: Counterfactual Economies without VC or Any Equity

expected growth trajectories of VC-backed firms under alternative sources of financing.[73] The counterfactual experiment predicts that firms born with the same growth potential would grow around 22% more in expectation with VC financing by age 10 compared to the expected firm size with Angel financing.[74] On the other hand, the expected firm size would decline by 53% with bank financing relative to Angel financing, demonstrating the importance of equity financing for high-growth firms.

At the macro level, a counterfactual economy is simulated in which VC financing is unavailable. VC-backed firms in the status quo would instead raise Angel financing in this scenario. This predicts about 0.4% smaller aggregate consumption in the steady state, as reported in the first column on Table 8, which is equivalent to around 70 billion dollars of total consumption expenditures in the US in 2022. Hence, I interpret it as a sizable impact on the economy.[75]

The second column shows the simulation results with an economy without any equity (VC or Angel) financing. This leads to around 0.8% lower aggregate consumption in the steady state. Equity-backed firms in the status quo now rely on bank financing and are less likely to seize their growth opportunities, leading to a loss in aggregate consumption.

---

[73] The average employment of surviving firms is multiplied by the survival rate to compare the expected firm size.

[74] In comparison to the literature, Sørensen (2007) estimates a two-sided matching model with heterogeneous firms and VCs and concludes that endogenous sorting is almost twice as important as the influence of VCs on the firm's performance.

[75] Personal consumption expenditures in US GDP in 2022 are around 17.5 trillion dollars (Source: US Bureau of Economic Analysis). As a reference, the US GDP declined by around 2.6% in 2009 during the Great Recession (Source: FRED). Lucas (1987)'s calculation of the cost of the business cycle is less than 0.1% of consumption, which is considered a small value. The value of 0.4% in this study is somewhat in between, so I interpret it as a medium-scale impact. Note that the model does not consider the positive spillover of innovation or its impact on economic growth; thus, the prediction can be seen as a lower bound of the role of VCs in the aggregate economy. In addition, the total labor supply is assumed to be constant regardless of changes in wages.

Finally, the third column reports the simulation results with more VCs. If the number of VCs in the economy doubles, more startups benefit from VCs' advice. In addition, more VCs in the economy decreases the equilibrium management fees of VCs, further facilitating the growth of VC-backed firms.

5 Concluding Remarks

This paper quantitatively examines the high growth of VC-backed firms in the US economy. Empirically, the newly constructed firm-level micro dataset is used to document the substantial growth of VC-backed firms at the micro level and their contribution to growth at the macro level. VC-backed firms are characterized by their large upfront investment and extra equity stakes acquired by VC investors. Based on the evidence, a firm dynamics model with innovation-based endogenous firm productivity and the choice of financing from VC, Angel, and bank financing is developed. The model explains the growth trajectories of VC-backed firms and implies that VCs' managerial advice accounts for around one-fourth of the growth of VC-backed firms. The importance of VC financing is the macro economy is quantified in a general equilibrium analysis.

Based on the results of this paper, several directions may be pursued. First, firm growth is analyzed in a stationary equilibrium in this paper. It will be beneficial to extend the model to an endogenous growth model in which the economy sustains long-run growth. Then, the impact of the VC sector on aggregate growth can be analyzed.

Second, the two roles of venture capitalists examined in this paper are equity-based funding and managerial advice. Another role of VCs is to screen startups and monitor the usage of funds by VC-backed firms. An extension to incorporate uncertainty in a firm's potential and information asymmetry between firm owners and investors will help us understand the impact of information frictions on firm growth and aggregate output.

References

Acemoglu, D., Akcigit, U., and Celik, M. A. (2022). Radical and incremental innovation: The roles of firms, managers, and innovators. *American Economic Journal: Macroeconomics*, 14(3):199–249.

Akcigit, U., Dinlersoz, E., Greenwood, J., and Penciakova, V. (2022). Synergizing ventures. *Journal of Economic Dynamics and Control*, page 104427.

Akcigit, U. and Kerr, W. R. (2018). Growth through heterogeneous innovations. *Journal of Political Economy*, 126(4):1374–1443.

Bernstein, S., Giroud, X., and Townsend, R. R. (2016). The impact of venture capital monitoring. *The Journal of Finance*, 71(4):1591–1622.

Bertoni, F., Colombo, M. G., and Grilli, L. (2011). Venture capital financing and the growth of high-tech start-ups: Disentangling treatment from selection effects. *Research Policy*, 40(7):1028–1043.

Boar, C., Gorea, D., and Midrigan, V. (2022). Why are returns to private business wealth so dispersed? Technical report, National Bureau of Economic Research.

Buera, F. J. and Shin, Y. (2013). Financial frictions and the persistence of history: A quantitative exploration. *Journal of Political Economy*, 121(2):221–272.

Catalini, C., Guzman, J., and Stern, S. (2019). Hidden in plain sight: Venture growth with or without venture capital. Technical report, National Bureau of Economic Research.

Chemmanur, T. J., Krishnan, K., and Nandy, D. K. (2011). How does venture capital financing improve efficiency in private firms? a look beneath the surface. *The Review of Financial Studies*, 24(12):4037–4090.

Chow, M. C., Fort, T. C., Goetz, C., Goldschlag, N., Lawrence, J., Perlman, E. R., Stinson, M., and White, T. K. (2021). Redesigning the longitudinal business database. Technical report, National Bureau of Economic Research.

Cohen, A. (2023). A Crosswalk from the Business R&D and Innovation Survey (BRDIS) and the Survey of Industrial Research and Development (SIRD) to the Longitudinal Business Database (LBD) and the Standard Statistical Es. CES Technical Notes Series 23-08, Center for Economic Studies, U.S. Census Bureau.

Cole, H. L., Greenwood, J., and Sanchez, J. M. (2016). Why doesn't technology flow from rich to poor countries? *Econometrica*, 84(4):1477–1521.

Cooley, T. F. and Quadrini, V. (2001). Financial markets and firm dynamics. *American Economic Review*, 91(5):1286–1310.

Corbae, D. and D'Erasmo, P. (2021). Reorganization or liquidation: Bankruptcy choice and firm dynamics. *The Review of Economic Studies*, 88(5):2239–2274.

Davis, S. J., Haltiwanger, J., and Schuh, S. (1996). Small business and job creation: Dissecting the myth and reassessing the facts. *Small Business Economics*, 8:297–315.

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

Decker, R. A., Haltiwanger, J., Jarmin, R. S., and Miranda, J. (2016). Where has all the skewness gone? the decline in high-growth (young) firms in the us. *European Economic Review*, 86:4–23.

Dent, R. C., Pugsley, B. W., and Wheeler, H. (2018). Longitudinal linking of enterprises in the lbd and ssl. Technical report, Center for Economic Studies, US Census Bureau.

Dinlersoz, E., Kalemli-Ozcan, S., Hyatt, H., and Penciakova, V. (2018). Leverage over the life cycle and implications for firm growth and shock responsiveness. Technical report, National Bureau of Economic Research.

Ewens, M., Gorbenko, A., and Korteweg, A. (2022). Venture capital contracts. *Journal of Financial Economics*, 143(1):131–158.

Ewens, M., Nanda, R., and Rhodes-Kropf, M. (2018). Cost of experimentation and the evolution of venture capital. *Journal of Financial Economics*, 128(3):422–442.

Feld, B. and Mendelson, J. (2019). *Venture deals: Be smarter than your lawyer and venture capitalist*. John Wiley & Sons.

Feld, B. and Ramsinghani, M. (2013). *Startup boards: Getting the most out of your board of directors*. John Wiley & Sons.

Fenn, G. W., Liang, N., and Prowse, S. (1997). The private equity market: An overveiw. *Financial Markets, Institutions & Instruments*, 6(4):1–106.

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

Fox, J. T., Hsu, D. H., and Yang, C. (2012). Unobserved heterogeneity in matching games with an application to venture capital. Technical report, National Bureau of Economic Research.

Gomes, J. F. (2001). Financing investment. *American Economic Review*, 91(5):1263–1285.

Gomes, J. F. and Sarkisyan, S. (2023). Monetary policy and financial stability. *Available at SSRN*.

Gompers, P. and Lerner, J. (2001). The venture capital revolution. *Journal of Economic Perspectives*, 15(2):145–168.

Gompers, P. A. (1994). The rise and fall of venture capital. *Business and Economic History*, 23(2):1–26.

Gompers, P. A., Gornall, W., Kaplan, S. N., and Strebulaev, I. A. (2020). How do venture capitalists make decisions? *Journal of Financial Economics*, 135(1):169–190.

Greenwood, J., Han, P., and Sanchez, J. M. (2022). Financing ventures. *International Economic Review*, 63(3):1021–1053.

Haltiwanger, J., Jarmin, R. S., Kulick, R., and Miranda, J. (2016). High growth young firms: contribution to job, output, and productivity growth. In *Measuring entrepreneurial businesses: Current knowledge and challenges*, pages 11–62. University of Chicago Press.

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

Hellmann, T., Schure, P., and Vo, D. H. (2021). Angels and venture capitalists: substitutes or complements? *Journal of Financial Economics*, 141(2):454–478.

Hennessy, C. A. and Whited, T. M. (2007). How costly is external financing? evidence from a structural estimation. *The Journal of Finance*, 62(4):1705–1745.

Hochberg, Y. V., Ljungqvist, A., and Lu, Y. (2007). Whom you know matters: Venture capital networks and investment performance. *The Journal of Finance*, 62(1):251–301.

Hopenhayn, H. A. (1992). Entry, exit, and firm dynamics in long run equilibrium. *Econometrica*, 60(5):1127–1150.

Hsu, D. H. (2004). What do entrepreneurs pay for venture capital affiliation? *The Journal of Finance*, 59(4):1805–1844.

Jarmin, R. S. and Miranda, J. (2002). The longitudinal business database. *Available at SSRN 2128793*.

Jovanovic, B. and Szentes, B. (2013). On the market for venture capital. *Journal of Political Economy*, 121(3):493–527.

Kenney, M. (2011). How venture capital became a component of the us national system of innovation. *Industrial and Corporate Change*, 20(6):1677–1723.

Khan, A., Senga, T., and Thomas, J. K. (2014). Default risk and aggregate fluctuations in an economy with production heterogeneity. *Unpublished manuscript*, 5.

Kochen, F. (2022). Finance over the life cycle of firms. *Unpublished manuscript*.

Kondo, I. O., Lewis, L. T., and Stella, A. (2023). Heavy tailed but not zipf: Firm and establishment size in the united states. *Journal of Applied Econometrics*.

Kortum, S. and Lerner, J. (2000). Assessing the contribution of venture capital to innovation. *RAND Journal of Economics*, 31(4):674–692.

Lerner, J. and Nanda, R. (2020). Venture capital's role in financing innovation: What we know and how much we still need to learn. *Journal of Economic Perspectives*, 34(3):237–261.

Lucas, R. E. (1987). *Models of business cycles*. B. Blackwell.

Luttmer, E. G. (2010). Models of growth and firm heterogeneity. *Annual Review of Economics*, 2(1):547–576.

Ma, Y. (2022). Specialization in a knowledge economy. *Available at SSRN 4052990*.

Metrick, A. and Yasuda, A. (2021). *Venture capital and the finance of innovation*. John Wiley & Sons.

Mezzanotti, F. and Simcoe, T. (2023). Innovation and appropriability: Revisiting the role of intellectual property. Technical report, National Bureau of Economic Research.

Midrigan, V. and Xu, D. Y. (2014). Finance and misallocation: Evidence from plant-level data. *American Economic Review*, 104(2):422–458.

Opp, C. C. (2019). Venture capital and the macroeconomy. *The Review of Financial Studies*, 32(11):4387–4446.

Ottonello, P. and Winberry, T. (2020). Financial heterogeneity and the investment channel of monetary policy. *Econometrica*, 88(6):2473–2502.

Ottonello, P. and Winberry, T. (2023). Investment, innovation, and financial frictions.

Pugsley, B. W. and Şahin, A. (2019). Grown-up business cycles. *The Review of Financial Studies*, 32(3):1102–1147.

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

Robb, A. M. and Robinson, D. T. (2014). The capital structure decisions of new firms. *The Review of Financial Studies*, 27(1):153–179.

Salgado, S. (2020). Technical change and entrepreneurship. *Available at SSRN 3616568*.

Salgado, S., Guvenen, F., and Bloom, N. (2019). Skewed business cycles. Technical report, National Bureau of Economic Research.

Sohl, J. (2012). The changing nature of the angel market. *The handbook of research on venture capital*, 2:17–41.

Sørensen, M. (2007). How smart is smart money? a two-sided matching model of venture capital. *The Journal of Finance*, 62(6):2725–2762.

Sterk, V., Sedláček, P., and Pugsley, B. (2021). The nature of firm growth. *American Economic Review*, 111(2):547–79.

Sui, X. (2020). Financial development, firm growth, and aggregate productivity divergence in europe. *Firm Growth, and Aggregate Productivity Divergence in Europe (June 30, 2020)*.

Wetzel Jr, W. E. (1983). Angels and informal risk capital. *Sloan Management Review (pre-1986)*, 24(4):23.

Online Appendix

A Appendix for the Empirical Results

A.1 Detailed Description of the Data

Longitudinal Business Database (LBD): The LBD contains records of all non-farm employer establishments that operate in the United States. The database is based on the administrative records from business tax filings that are shared by the Internal Revenue Service (IRS) with the US Census Bureau. The Census Bureau restores the longitudinal link of establishments across time based on tax identifiers and establishment names and addresses. See Chow et al. (2021) for the Census Bureau's recent redesign of the LBD.

A challenge is to construct longitudinally consistent firm identifiers. A firm is an economic unit consisting of single or multiple establishments under common ownership. The LBD historically used `firmid`, which is "0" followed by Employer Identification Number (`EIN`) for single-unit firms and `alpha` followed by "0000" for multi-unit firms. The `firmid` uniquely identifies all firms in each year, but the longitudinal linkage is not complete. This is because `firmid` may change over time for reasons unrelated to ownership changes.

I partially restore firm links over time for two cases.[76] First, if a firm transitions from a single-unit to a multi-unit firm (in this case, `firmid` always changes by construction) while preserving its `EIN`, then I treat the two firms with different `firmid` as the same firm. Second, a firm obtains a new `EIN` when its legal form of organization (LFO) transforms. Therefore, if a firm changes its LFO while preserving its `lbdnum`, a longitudinal establishment identifier created in the Census Bureau, and single-unit status, I treat the two firms as the same firm. These two efforts are particularly important for keeping track of high-growth firms and VC-backed firms as they are likely to change their single-unit/multi-unit status and LFO over the life cycle. These efforts also reduce the number of spurious firm births and deaths.

Survey of Business Owners (2007): The Survey of Business Owners (SBO), conducted by the Census Bureau, asks a large sample of firms in the US about their business and owner characteristics. Around 62% of the 2.3 million businesses in the survey in 2007 responded. The fraction of each group of firms is tabulated in Table 9, using sampling weights in the dataset.

The businesses in the SBO (2007) that responded to the questionnaire about the sources of capital are linked to the LBD. VC-backed firms are identified based on the name and address matching with the external VC datasets. Among the firms that received an early-stage investment in exchange for their equity, according to the questionnaire in the SBO, those that did not receive VC investment are regarded as non-VC-equity-

---

[76] I referred to the Center for Economic Studies (CES) technical notes by Dent et al. (2018) to implement this analysis. They have worked on creating firm-level longitudinal identifiers that are robust to ownership changes.

	Fraction in the economy	Odds of being Gazelles
VC-backed firms	0.20%	46.5%
Non-VC equity	0.75%	10.8%
Bank-financed	31%	2.57%

Notes: The businesses in the SBO (2007) that replied to the questionnaire about the sources of capital are linked to the LBD. Payroll Gazelles are firms with an average annual growth rate of payroll above 20% in the first five years and whose payroll reaches 100 or above multiplied by the average payroll in the economy by age 20. Observations are weighted by the sampling weights in the SBO.

Table 9: Firms in SBO (2007) and payroll Gazelles

financed firms. Bank-backed firms are those receiving a "business loan from a bank or financial institution" or "government-guaranteed business loan from a bank or financial institution" but not obtaining VC financing or non-VC equity financing.[77]

Table 9 reports the fraction of payroll Gazelles in the three groups and shows that VC-backed firms are more likely to be Gazelles (46.5%) than non-VC-equity-backed firms (10.8%) and bank-backed firms (2.57%). In contrast to the fraction of Gazelles among newborn firms documented in Table 2, this table shows the fraction of Gazelles among firms operated in 2007. The fraction of Gazelles among VC-backed firms in this sample is higher, as VC-backed firms that did not achieve high growth are more likely to have exited. As we have seen in the right panel in Figure 1, VC-backed firms keep their substantial growth over the life cycle relative to non-VC-equity-financed firms or bank-financed firms conditional on being high-growth firms.

SIRD (1980-2007) and BRDIS (2008-2019): The Survey of Industrial Research and Development (SIRD) from 1980 to 2007 and the Business R&D and Innovation Survey (BRDIS) from 2008 to 2019 provide information on the R&D activity of US-based companies. The National Science Foundation (NSF) sponsors the surveys. Approximately 25,000–40,000 companies are surveyed each year. The survey's name changed in 2007, but there is no break in the series of items collected in both surveys. The key variable analyzed in this paper is the total costs incurred for R&D. The two datasets are linked to

---

[77] SBO (2007) asks about the sources of capital. The choices are as follows: Personal/family savings of owner(s), Personal/family assets other than savings of owner(s), Personal/family home equity loan, Personal/business credit card(s), Business loan from federal, state, or local government, Government-guaranteed business loan from a bank or financial institution, Business loan from a bank or financial institution, Business loan/investment from family/friend(s), Investment by venture capitalist(s) (an early-stage investment in exchange for ownership equity by an individual, outside group, or business not directly involved in the overall operation and management of the business), Grants, Other source(s) of capital, Don't know, and None needed.

the LBD using census identifiers.[78]

A.2 Additional Empirical Results

High Growth of VC-backed Gazelles: Figure 1 in Section 2 shows the growth trajectories of Gazelles measured in the firm's average payroll. In this appendix, the robustness of the pattern is shown. Figure 13 compares the growth trajectories of VC-backed and non-VC-backed Gazelles in terms of average employment. Figure 14 shows the growth trajectories in terms of the (pseudo) median payroll. Finally, Figure 15 plots regression coefficients on age interacted with VC-backing status in a Deaton-Paxson regression with firm revenue as the dependent variable and year and NAICS (four digits) as the control variables. All figures show that VC-backed Gazelles attain high growth compared to non-VC-backed Gazelles.

Table 10 shows the regression results for the growth rate of firms in the sample of Gazelles. This regression addresses two issues in Figure 1. In the figure, the average firm size is computed among surviving firms. Also, the firm size may increase by acquiring another firm. Here, since the dependent variable measures the DHS growth rate, the growth rate is -2 if the firm exits in the next period instead of dropping from the sample. Additionally, the growth rate is robust to ownership changes since the firm growth is measured as the sum of the organic expansion of establishments that belong to the firm in the next period. After taking into account the two potential concerns (survival bias and ownership changes), VC-backed firms are shown to attain high growth compared to non-VC-backed firms. Note that after controlling for firm age, the coefficients on firm size measured by log(emp) and log(payroll) are statistically insignificant. This finding is consistent with the finding by Haltiwanger et al. (2013) that after controlling for firm age, firm growth is not systematically related to firm size.

Finally, Table 11 compares the exit rate between VC-backed firms and non-VC-backed firms. Among firms aged 0 to 20, the exit rate of VC-backed firms is around 2.5% lower compared to non-VC-backed firms after controlling for year, age, and NAICS (four digits). In the sample of Gazelles, the difference in exit rates is statistically insignificant in the sample of employment Gazelles and is positive with around 1.2% in the sample of payroll Gazelles.

B Mathematical Derivations

B.1 Equity Share

This section derives the equivalence of two formulations, maximizing the firm's total value and maximizing the startup founder's value, along with the evolution of equity

---

[78] The Center for Economic Studies (CES) technical notes by Cohen (2023) are referred to implement it.

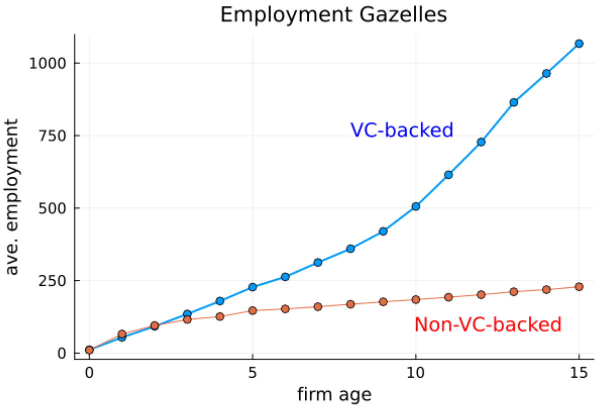


Figure 13: Average Employment

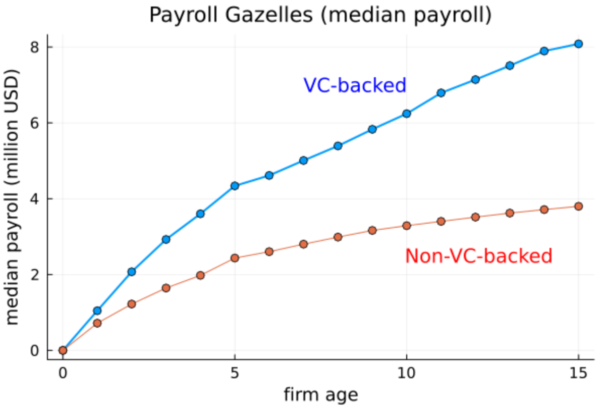
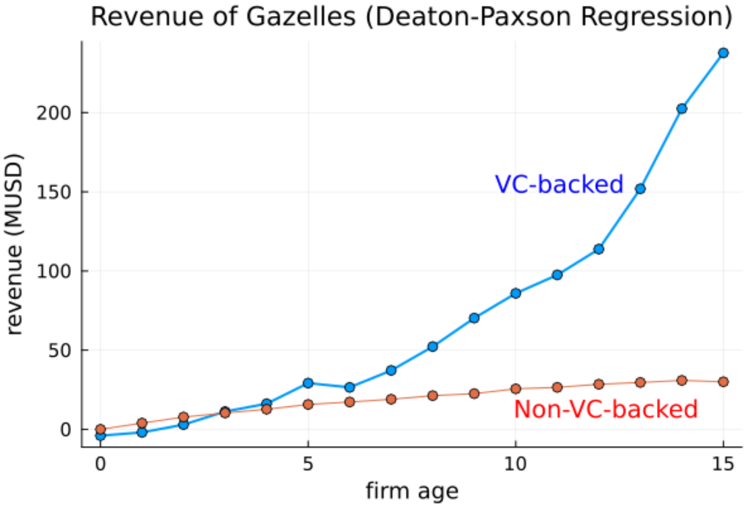


Figure 14: Median Payroll



Notes: In the Deaton-Paxson regression, the dependent variable is firm revenue. The independent variables are firm age interacted with VC-backed dummy, year, and NAICS four-digit codes. The coefficients on the age dummies are plotted in the figure. Data: LBD (1997-2019) and VC datasets.

Figure 15: Deaton-Paxson Regression for firm's revenue

	Dependent variable (%):	
	(1) emp growth	(2) payroll growth
VC-backed	3.76 ***	3.02***
	(0.691)	(0.705)
firm age	–1.39***	–1.31***
	(0.0571)	(0.152)
log(emp)	–0.0603	
	(0.246)	
log(payroll)		0.251
		(0.246)
Year	Yes	Yes
Initial Year	Yes	Yes
NAICS (4 digits)	Yes	Yes
N	2,070,000	2,235,000
R^2	0.036	0.046

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Notes: The sample of Gazelles, employment Gazelles in (1) and payroll Gazelles in (2), born from 1980 to 2010 at ages 0 to 20. Gazelles are firms with an average annual growth rate above 20% in the first five years and employment of 100 or above by age 20. The dependent variables are the DHS growth rates of employment and payrolls, $g_{it} = \frac{E_{i,t} - E_{i,t-1}}{0.5 * (E_{i,t} + E_{i,t-1})}$. Data: LBD (1980-2019) and VC datasets.

Table 10: Growth rate of Gazelles (1980-2019)

	Dependent variable: exit at $t+1$		
	(1) all firms	(2) emp Gazelles	(3) payroll Gazelles
VC-backed	–0.0248***	0.00104	0.0120***
	(0.000491)	(0.000643)	(0.000585)
Year	Yes	Yes	Yes
Age Dummy	Yes	Yes	Yes
NAICS (4 digits)	Yes	Yes	Yes
N	136,200,000	2,185,000	2,348,000
R^2	0.015	0.013	0.011

Standard errors in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Notes: LHS is 1 if a firm exits in the next period. This is the sample of firms born from 1980 to 2010 at ages 0 to 20. Data: LBD (1980-2019) and VC datasets.

Table 11: Comparison of Exit Rate between VC-backed firms and Non-VC-backed firms

stakes owned by investors. This equivalence is an intuitive result, as investors and startup founders are both risk-neutral in the model.

I consider the case $d_t < 0$, where the firm raises equity and distributes zero dividends. The amount of money raised is denoted by $m_t = -d_t$. I rewrite the firm's maximization problem in equity financing in equation (9). For simplicity, I assume no exogenous exit shock in this derivation:

$$V_t^{\text{total}} = \max_{m_t} -(1+c_{e,1})m_t + \frac{1}{1+r}\mathbb{E}\left[V_{t+1}^{\text{total}}\right] \tag{37}$$

$$\text{s.t. expend} = \text{rev} + m_t, \tag{38}$$

where expend $= Rk_{t+1} + w\ell_{t+1} + I_t + c_f + c_{e,0} + \mathbf{1}\{\text{VC}\}\bar{\pi}_{\text{vc}}$ and rev $= z^{1-\eta}(k^{\alpha}\ell^{1-\alpha})^{\eta}$. The investors inject m_t amount of money into the firm, which costs $(1+c_{e,1})m_t$ as a result of the flotation costs $c_{e,1}m_t$. The discounted firm value after the investment, which is called post-money valuation, is given by $\frac{1}{1+r}\mathbb{E}\left[V_{t+1}^{\text{total}}\right]$.

On the other hand, the startup founder's value is formulated as follows. The founder holds $(1-s_t)$ fraction of the firm's pre-money valuation, where $V_t^{\text{pre}} = V_t^{\text{post}} - X$ and $X = (1+c_{e,1})m_t$. The pre-money valuation corresponds to V_t^{total} in the formulation above.

The founder's maximization problem is given by:

$$V_t^{\text{startup}} = \max_{m_t}(1 - s_{t+1})\frac{1}{1+r}\mathbb{E}\left[V_{t+1}^{\text{pre}}\right] \tag{39}$$

$$\text{s.t. } s_{t+1} = s_t + (1 - s_t)\frac{X}{V_t^{\text{post}}} \tag{40}$$

$$\text{expend} = \text{rev} + m_t. \tag{41}$$

The founder owns $(1 - s_{t+1})$ fraction of the firm ownership after the financial investment at t. I first verify the evolution of equity stakes in equation (40) using the investor's zero-profit conditions. Investors at the beginning of the period hold s_t fraction of the firm's pre-money valuation. The value of their equity stakes should be the same after raising additional equity this period and is equivalent to the $\tilde{s}_t$ fraction of post-money valuation:

$$s_t V_t^{\text{pre}} = \tilde{s}_t V_t^{\text{post}}. \tag{42}$$

New equity investments inject m_t by paying $X = (1 + c_{e,1})m_t$ in exchange for the $\tilde{s}_{t+1}$ fraction of the firm's post-money valuation. Their zero-profit condition implies:

$$X = \tilde{s}_{t+1} V_t^{\text{post}}. \tag{43}$$

Hence, the total equity shares held by the investors are:

$$\begin{aligned} s_{t+1} &= \tilde{s}_t + \tilde{s}_{t+1} \\ &= s_t \frac{V_t^{\text{pre}}}{V_t^{\text{post}}} + \frac{X}{V_t^{\text{post}}} \\ &= s_t + (1 - s_t)\frac{X}{V_t^{\text{post}}}. \end{aligned}$$

Hence, equation (40) is obtained. Now, I substitute (40) into (39). Since the founder's equity stakes in the next period are given by:

$$1 - s_{t+1} = (1 - s_t)\left(1 - \frac{X}{V_t^{\text{post}}}\right),$$

the founder's value at time t is given by:

$$V_t^{\text{startup}} = \max_{m_t}(1 - s_t)\left[1 - \frac{X}{V_t^{\text{post}}}\right]\frac{1}{1+r}\mathbb{E}\left[V_{t+1}^{\text{pre}}\right]. \tag{44}$$

Since the firm distributes zero dividends in this period, the post-money valuation is:

$$V_t^{\text{post}} = \frac{1}{1+r}\mathbb{E}\left[V_{t+1}^{\text{pre}}\right]. \tag{45}$$

Hence, the objective function (39) is written as:

$$V_t^{\text{startup}} = (1 - s_t)\left\{\max_{m_t} -(1 + c_{e,1})m_t + \frac{1}{1+r}\mathbb{E}\left[V_{t+1}^{\text{pre}}\right]\right\}$$

$$\text{s.t. expend} = \text{rev} + m_t. \tag{46}$$

It can be seen that maximizing the founder's value, which is $(1 - s_t)$ fraction of the firm's value, is equivalent to maximizing the firm's total value in (37).

B.2 Optimality Conditions

B.2.1 Economy without Financial Frictions or Innovation Expenditures

I describe optimality conditions in a frictionless economy without innovation expenditures in this subsection. Although it is not directly relevant to this paper, the optimality conditions are simple and easy to interpret without financial frictions and innovation. In particular, if the firm's productivity in the next period (z') is fixed, then firm size and firm profits are linear in z'.

Optimal labor and capital (ℓ', k') are chosen to equalize the marginal benefits and marginal costs of the inputs. The maximization problem is given by:

$$V = \max_{\ell', k'} y - w\ell' - Rk' - c_f + \frac{1}{1+r}\mathbb{E}\left[(1-\varphi)V' + \varphi y'\right]$$

where $y = z^{1-\eta}(k^{\alpha}\ell^{1-\alpha})^{\eta}$.

FOCs are:

$$\begin{aligned} w &= \frac{1}{1+r}\mathbb{E}\left[y'\frac{(1-\alpha)\eta}{\ell'}\right] \\ R &= \frac{1}{1+r}\mathbb{E}\left[y'\frac{\alpha\eta}{k'}\right]. \end{aligned}$$

This gives an optimality condition between ℓ' and k':

$$\ell' = \frac{1-\alpha}{w}\frac{R}{\alpha}k'. \tag{47}$$

The optimal (ℓ', k') are:

$$\begin{aligned} \ell' &= \mathbb{E}[(z')^{1-\eta}]^{\frac{1}{1-\eta}}\left\{\frac{\eta}{1+r}\left(\frac{\alpha}{R}\right)^{\alpha\eta}\left(\frac{1-\alpha}{w}\right)^{1-\alpha\eta}\right\}^{\frac{1}{1-\eta}} \\ k' &= \mathbb{E}[(z')^{1-\eta}]^{\frac{1}{1-\eta}}\left\{\frac{\eta}{1+r}\left(\frac{\alpha}{R}\right)^{1-(1-\alpha)\eta}\left(\frac{1-\alpha}{w}\right)^{(1-\alpha)\eta}\right\}^{\frac{1}{1-\eta}}. \end{aligned}$$

The expected output and expected profits are given as follows:

$$\begin{aligned} \mathbb{E}[y'] &= \mathbb{E}[(z')^{1-\eta}]^{\frac{1}{1-\eta}}\left[\frac{\eta}{1+r}\left(\frac{\alpha}{R}\right)^{\alpha}\left(\frac{1-\alpha}{w}\right)^{1-\alpha}\right]^{\frac{\eta}{1-\eta}} \\ \mathbb{E}[\pi'] &:= \frac{1}{1+r}\mathbb{E}[y'] - w\ell' - Rk' - c_f = \frac{\mathbb{E}[y']}{1+r} - \frac{R}{\alpha}k' - c_f \\ &= \mathbb{E}[(z')^{1-\eta}]^{\frac{1}{1-\eta}}\left[\eta^{\frac{\eta}{1-\eta}} - \eta^{\frac{1}{1-\eta}}\right]\left(\frac{1}{1+r}\right)^{\frac{1}{1-\eta}}\left[\left(\frac{\alpha}{R}\right)^{\alpha}\left(\frac{1-\alpha}{w}\right)^{1-\alpha}\right]^{\frac{\eta}{1-\eta}} \\ &= \mathbb{E}[(z')^{1-\eta}]^{\frac{1}{1-\eta}}(1-\eta)\left(\frac{1}{1+r}\right)^{\frac{1}{1-\eta}}\left[\left(\frac{\alpha\eta}{R}\right)^{\alpha\eta}\left(\frac{(1-\alpha)\eta}{w}\right)^{(1-\alpha)\eta}\right]^{\frac{1}{1-\eta}}. \end{aligned}$$

If productivity in the next period (z') is fixed, labor and capital inputs (ℓ', k'), output in the next period (y'), and profits in the next period π' are all linear in z'.

B.2.2 Economy without Financial Frictions

I derive the optimality conditions with respect to capital inputs and innovation expenditures in the economy without financial frictions but with innovation expenditures. The value function in this setup can be solved very quickly, so I calibrate this economy first to obtain a reasonable initial guess of parameters in the economy with financial frictions.[79]

Firms are born with growth potential $\hat{\theta}$ with corresponding productivity ladders $(\hat{\theta}_1, \cdots, \hat{\theta}_N)$. Denote the current position in the productivity ladder by n. The value function is given by the following. I substitute the optimal labor inputs as a function of capital inputs, as derived in (47). Since firms do not face financial frictions, the cost of raising one unit of capital is one.

$$\begin{aligned} V(n, \epsilon, k, \text{fail}; \hat{\theta}) = y + & \\ + \max \Bigg\{ 0, \max_{k', I} & -\frac{R}{\alpha}k' - I - c_f + \frac{1-\varphi}{1+r}\mathbb{E}_{n',\epsilon'}\left[V(n', \epsilon', k', \text{fail}; \hat{\theta}) | \varsigma, \epsilon\right] \\ & + \frac{\varphi}{1+r}E_{n',\epsilon'}\left[y(n', \epsilon', k'; \hat{\theta}) | \varsigma, \epsilon\right] \Bigg\} \end{aligned} \tag{48}$$

$$\begin{aligned} \text{where } y &= (\hat{\theta}_{(n)} + \epsilon)(k^{\alpha} l^{1-\alpha})^{\eta} \\ &= (\hat{\theta}_{(n)} + \epsilon)\left(\frac{1-\alpha}{\alpha}\frac{R}{w}\right)^{(1-\alpha)\eta} k^{\eta} \end{aligned} \tag{49}$$

The FOC with respect to k' is:

$$\underbrace{\frac{R}{\alpha}}_{\text{MC}} = \frac{1-\varphi}{1+r}\mathbb{E}_{n',\epsilon'}\left[V_{k'}(n', \epsilon', \text{fail}, k', \text{fail}; \hat{\theta}) | \varsigma, \epsilon\right] + \frac{\varphi}{1+r}E_{n',\epsilon'}\left[\frac{\partial}{\partial k'}y(n', \epsilon', k'; \hat{\theta}) | \varsigma, \epsilon\right],$$

$$\text{where } V_k(n, \epsilon, k, \text{fail}; \hat{\theta}) = \frac{\partial}{\partial k}y(n, \epsilon, k; \hat{\theta}) = (\hat{\theta}_{(n)} + \epsilon)\left(\frac{1-\alpha}{\alpha}\frac{R}{w}\right)^{(1-\alpha)\eta} \eta k^{\eta-1}.$$

The optimal k' is derived as:

$$k' = \left\{\frac{\alpha}{R}\left(\frac{1-\alpha}{\alpha}\frac{R}{w}\right)^{(1-\alpha)\eta} \frac{\eta}{1+r}\left(\mathbb{E}_{n'}[\theta_{n'} | \varsigma^*] + \mathbb{E}_{\epsilon'}\left[\epsilon' | \epsilon\right]\right)\right\}^{\frac{1}{1-\eta}}. \tag{50}$$

The FOC with respect to I is:

$$\begin{aligned} 1 = & \frac{1-\varphi}{1+r}\frac{\partial \varsigma}{\partial I}\mathbb{E}_{\epsilon'}\left[V(n+1, \epsilon', k', \text{fail} = 0; \hat{\theta}) - V(n, \epsilon', k', \text{fail} = 1; \hat{\theta}) | \epsilon\right] \\ & + \frac{\varphi}{1+r}\frac{\partial \varsigma}{\partial I}\mathbb{E}_{\epsilon'}\left[y(n+1, \epsilon', k'; \hat{\theta}) - y(n, \epsilon', k'; \hat{\theta}) | \epsilon\right]. \end{aligned} \tag{51}$$

---

[79] Given the parameters, it takes less than one second to compute the stationary equilibrium in this economy.

Since $\frac{\partial \varsigma}{\partial I} = \varsigma(1-\varsigma)\frac{\gamma_1}{I}$, the optimal I follows:

$$\underbrace{1}_{\text{MC}} = \frac{\gamma_1 \varsigma^*(1-\varsigma^*)}{(1+r)I^*}\Bigg\{(1-\varphi)\mathbb{E}_{\epsilon'}\left[V(n+1,\epsilon',k',\text{fail}=0;\hat{\theta}) - V(n,\epsilon',k',\text{fail}=1;\hat{\theta})\Big|\epsilon\right] \\ + \varphi(\hat{\theta}_{n+1}-\hat{\theta}_n)\left(\frac{1-\alpha}{\alpha}\frac{R}{w}\right)^{(1-\alpha)\eta}(k')^{\eta}\Bigg\}. \tag{52}$$

The optimal (k', I) is a solution to a system of two equations (50) and (52) and is computed with a non-linear solver.

B.2.3 Equity Financing

The value function of equity-backed firms is formulated in Section 3.2. The Lagrangian formulation of the value function of equity-financed firms that have decided to continue the business is as follows:

$$\begin{aligned} V^{j,C}(n,\epsilon,o,\text{fail};\hat{\theta}) = & \max_{d,k',I} d + \mathbf{1}\{d<0\}(dc_{e,1}) \\ & + \frac{1-\varphi}{1+r}\mathbb{E}_{n',\epsilon'}\left[V^j(n',\epsilon',o',\text{fail}';\hat{\theta})|\varsigma,\epsilon\right] + \frac{\varphi}{1+r}\mathbb{E}_{n',\epsilon'}\left[o'|\varsigma,\epsilon\right] \\ & + \mu\left[o - \frac{R}{\alpha}k' - I - c_f - c_{e,0} - \mathbf{1}\{\text{VC}\}\bar{\pi}_{\text{vc}} - d\right] \end{aligned} \tag{53}$$

$$\text{s.t. } o = (\hat{\theta}_n+\epsilon)^{1-\eta}\left(\frac{1-\alpha}{\alpha}\frac{R}{w}\right)^{(1-\alpha)\eta}k^{\eta}. \tag{54}$$

The FOCs with respect to the three choice variables are:

$$[d]: \quad (1+\mathbf{1}\{d<0\}c_{e,1}-\mu)\,d = 0 \tag{55}$$

$$[k']: \quad \frac{1-\varphi}{1+r}\mathbb{E}_{n',\epsilon'}\left[\frac{\partial V'}{\partial k'}\Big|\varsigma,\epsilon\right] + \frac{\varphi}{1+r}\mathbb{E}_{n',\epsilon'}\left[\frac{\partial y'}{\partial k'}\Big|\varsigma,\epsilon\right] = \underbrace{\frac{R}{\alpha}\mu}_{\text{MC}} \tag{56}$$

$$[I]: \quad \frac{\partial\varsigma}{\partial I}\left\{\frac{1-\varphi}{1+r}\,\mathbb{E}_{\epsilon'}\left[V(n+1)-V(n)\right] + \frac{\varphi}{1+r}\mathbb{E}_{\epsilon'}\left[o'(n+1)-o'(n)\right]\right\} = \underbrace{\mu}_{\text{MC}}. \tag{57}$$

The marginal cost of raising additional financial capital is given by μ, instead of 1, in equity financing. If $d > 0$, since the firm is not liquidity constrained, $\mu = 1$. If $d < 0$, since the firm relies on equity financing, $\mu = 1+c_{e,1}$. If $d = 0$, the Lagrangian multiplier μ satisfies $\mu \geq 1$, (56), and (57), while (k', I) follows the budget constraint. The derivative of the value function is given by:

$$\frac{\partial V}{\partial k} = \mu\frac{\partial y}{\partial k}. \tag{58}$$

By substituting:

$$\begin{aligned} \frac{\partial y}{\partial k} &= (\hat{\theta}_n+\epsilon)^{1-\eta}\left(\frac{1-\alpha}{\alpha}\frac{R}{w}\right)^{(1-\alpha)\eta}\eta k^{\eta-1} \\ \frac{\partial\varsigma}{\partial I} &= \varsigma^*(1-\varsigma^*)\frac{\gamma_1}{I}, \end{aligned}$$

the conditions are summarized as follows:

$$[d]: \quad \mu \begin{cases} = 1 & \text{if } d > 0 \\ \geq 1 & \text{if } d = 0 \\ = 1 + c_{e,1} & \text{if } d < 0 \end{cases} \tag{59}$$

$$[k']: \quad \left\{ \frac{1-\varphi}{1+r} \mathbb{E}_{n',\epsilon'} \left[(\hat{\theta}_{n'} + \epsilon')^{1-\eta} \mu' \right] + \frac{\varphi}{1+r} \mathbb{E}_{n',\epsilon'} \left[(\hat{\theta}_{n'} + \epsilon')^{1-\eta} \right] \right\} \left(\frac{1-\alpha}{\alpha} \frac{R}{w} \right)^{(1-\alpha)\eta} \eta k^{\eta-1} = \frac{R}{\alpha} \mu \tag{60}$$

$$[I]: \quad \frac{\gamma_1 \varsigma^* (1 - \varsigma^*)}{I} \left\{ \frac{1-\varphi}{1+r} \mathbb{E}_{\epsilon'} \left[V(n+1) - V(n) \right] + \frac{\varphi}{1+r} \mathbb{E}_{\epsilon'} \left[o'(n+1) - o'(n) \right] \right\} = \mu. \tag{61}$$

To numerically solve this problem, the optimal choice of (k', I, d) is computed for each case $(d > 0, d = 0, d < 0)$. The choice satisfying (59)–(61) and the budget constraint and maximizing the continuation value is chosen as a solution. The value function is iterated until it converges.

B.2.4 Bank Financing

The value function in bank financing is formulated in Section 3.3. The continuation value is given by:

$$V_t^{\text{Bank},C}(n_t, \epsilon_t, o_t, \text{fail}_t; \hat{\theta}) = \max_{d_t \geq 0, \ell_{t+1}, k_{t+1}, I_t, b_{t+1}} d_t \tag{62}$$

$$+ \frac{1}{1+r} \left\{ (1-\varphi) \underbrace{\mathbb{E}_{n_{t+1},\epsilon_{t+1}} \left[V_{t+1}(n_{t+1}, \epsilon_{t+1}, o_{t+1}, \text{fail}_{t+1}; \hat{\theta}) | \varsigma_{t+1} \right]}_{\text{no exit shock}} + \varphi \underbrace{\mathbb{E}_{n_{t+1},\epsilon_{t+1}} \left[\max\{0, o_{t+1}\} | \varsigma_{t+1} \right]}_{\text{exogenous exit}} \right\}$$

$$\text{s.t. } d_t + \underbrace{R k_{t+1} + w \ell_{t+1} + I_t + c_f}_{\text{expenditures}} = \underbrace{o_t}_{\text{cash flow}} + \underbrace{q_t(b_{t+1}, k_{t+1}, I_t) b_{t+1}}_{\text{new debt}} \tag{63}$$

$$o_t = (\hat{\theta}_{n_t} + \epsilon_t)^{1-\eta} (k_t^{\alpha} \ell_t^{1-\alpha})^{\eta} - b_t. \tag{64}$$

Equation (63) is the budget constraint, and equation (64) defines cash flow as revenue minus debt repayment. The price of debt, $q_t(b_{t+1}, k_{t+1}, I_t; n_t, \epsilon_t, o_t, \text{fail}_t, \hat{\theta})$, follows a zero-expected-profit condition of banks:

$$q_t(\cdot) = \frac{1}{1+r} \mathbb{E}_t \left[\underbrace{\mathbf{1}\{\Delta_{t+1} = 0\}}_{\text{no default}} + \underbrace{\mathbf{1}\{\Delta_{t+1} = 1\}}_{\text{default}} \left(\frac{y_{t+1}}{b_{t+1}} - c_d \right) \right]. \tag{65}$$

The term $\left(\frac{y_{t+1}}{b_{t+1}} - c_d \right)$ is a return of debt for banks in case of default.

Before describing the value function iteration, two properties can be noticed. First, the FOCs with respect to (ℓ_{t+1}, k_{t+1}) results in the following:[80]

$$\ell_{t+1} = \frac{(1-\alpha)R}{\alpha w} k_{t+1}. \tag{66}$$

Thus, one control variable (ℓ_{t+1}) is eliminated from the value function iteration. Second, by substituting d_t and $q_t(\cdot)$ using equations (63) and (65), the continuation value can be expressed without the price of debt (q_t) as follows:[81]

$$\begin{aligned}
V_t^{\text{Bank,C}}(n_t, \epsilon_t, o_t, \text{fail}_t; \hat{\theta}) &= \max_{k_{t+1}, I_t, b_{t+1}} o_t - \frac{R}{\alpha} k_{t+1} - I_t - c_f \\
&+ \frac{1}{1+r} \mathbb{E}_t \Bigg[(1-\varphi) \Big\{ \mathbf{1}\left\{V_{t+1}^{\text{bank}}(\cdot) \geq 0\right\} \left(V_{t+1}^{\text{bank}}(n_{t+1}, \epsilon_{t+1}, o_{t+1}, \text{fail}_{t+1}) + b_{t+1}\right) \\
&+ \mathbf{1}\left\{V_{t+1}^{\text{bank}}(\cdot) < 0 \wedge o_{t+1} \geq 0\right\} y_{t+1} + \left\{V_{t+1}^{\text{bank}}(\cdot) < 0 \wedge o_{t+1} < 0\right\} (y_{t+1} - c_d b_{t+1}) \Big\} \\
&+ \varphi \Big\{ \mathbf{1}\left\{o_{t+1} \geq 0\right\} y_{t+1} + \mathbf{1}\left\{o_{t+1} < 0\right\} (y_{t+1} - c_d b_{t+1}) \Big\} \Bigg],
\end{aligned} \tag{67}$$

$$\text{where } y_{t+1} = (\hat{\theta}_{n_{t+1}} + \epsilon_{t+1})^{1-\eta} \left(\frac{1-\alpha}{\alpha} \frac{R}{w} \right)^{(1-\alpha)\eta} k_{t+1}^{\eta}. \tag{68}$$

Therefore, the continuation value $V_t^{\text{Bank,C}}(\cdot)$ can be computed without separately computing the pricing function $q_t(\cdot)$.

---

[80] The FOCs with respect to ℓ_{t+1} and k_{t+1} give:

$$\begin{aligned}
\frac{1}{1+r} \mathbb{E} \left[\frac{\partial V_{t+1}}{\partial o_{t+1}} \frac{\partial o_t}{\partial \ell_{t+1}} + \mu \frac{\partial q_t}{\partial o_{t+1}} \frac{\partial o_t}{\partial \ell_{t+1}} b_{t+1} \right] &= \mu w \\
\frac{1}{1+r} \mathbb{E} \left[\frac{\partial V_{t+1}}{\partial o_{t+1}} \frac{\partial o_t}{\partial k_{t+1}} + \mu \frac{\partial q_t}{\partial o_{t+1}} \frac{\partial o_t}{\partial k_{t+1}} b_{t+1} \right] &= \mu R
\end{aligned}$$

with

$$\frac{\partial o_t}{\partial \ell_{t+1}} = (1-\alpha) \frac{y_{t+1}}{\ell_{t+1}} \quad \text{and} \quad \frac{\partial o_t}{\partial k_{t+1}} = \alpha \frac{y_{t+1}}{k_{t+1}}.$$

[81] The derivation in a case without an exogenous shock is as follows. This is done by substituting dividend (d_t) using a budget constraint.

$$\begin{aligned}
V_t^{\text{bank}}(n_t, \epsilon_t, o_t, \text{fail}_t; \hat{\theta}) &= \max_{k_{t+1}, I_t, b_{t+1}} o_t \\
&+ \frac{1}{1+r} \mathbb{E}_t \left[\mathbf{1}\left\{\Delta_{t+1} = 0\right\} + \mathbf{1}\left\{\Delta_{t+1} = 1\right\} \left(\frac{y_{t+1}}{b_{t+1}} - c_d \right) \right] b_{t+1} - \frac{R}{\alpha} k_{t+1} - I_t - c_f \\
&+ \frac{1}{1+r} \mathbb{E}_t \left[\mathbf{1}\left\{\text{continue}_{t+1}\right\} V_t^{\text{bank}}(n_{t+1}, \epsilon_{t+1}, o_{t+1}, \text{fail}_{t+1}) + \mathbf{1}\left\{\text{exit}_{t+1}\right\} o_{t+1} + \mathbf{1}\left\{\Delta_{t+1} = 1\right\} 0 \right] \\
&= \max_{k_{t+1}, I_t, b_{t+1}} o_t - \frac{R}{\alpha} k_{t+1} - I_t - c_f \\
&+ \frac{1}{1+r} \mathbb{E}_t \Big[\mathbf{1}\left\{\text{cont}_{t+1}(n_{t+1}, \epsilon_{t+1}, o_{t+1}, \text{fail}_{t+1})\right\} \left(V_{t+1}^{\text{bank}}(n_{t+1}, \epsilon_{t+1}, o_{t+1}, \text{fail}_{t+1}) + b_{t+1}\right) \\
&+ \mathbf{1}\left\{\text{exit}_{t+1}(n_{t+1}, \epsilon_{t+1}, o_{t+1}, \text{fail}_{t+1})\right\} y_{t+1} + \mathbf{1}\left\{\Delta_{t+1}(n_{t+1}, \epsilon_{t+1}, o_{t+1}, \text{fail}_{t+1}) = 1\right\} (y_{t+1} - c_d b_{t+1}) \Big], \\
&\text{where } y_{t+1} = (\hat{\theta}_{n_{t+1}} + \epsilon_{t+1})^{1-\eta} \left(\frac{1-\alpha}{\alpha} \frac{R}{w} \right)^{(1-\alpha)\eta} k_{t+1}^{\eta}.
\end{aligned}$$

Lagrangian Formulation The value function is computed based on the Lagrangian formulation:

$$V_t^{\text{Bank},C}(o_t, \epsilon_t, \text{fail}_t, n_t; \hat{\theta}) = \max_{d_t \geq 0, k_{t+1}, I_t, b_{t+1}} d_t + \frac{1}{1+r}\mathbb{E}\left[V_{t+1}^{\text{bank}}(o_{t+1}, \epsilon_{t+1}, \text{fail}_{t+1}, n_{t+1}; \hat{\theta})\right]$$
$$+ \mu_t \left[o_t + q_t(b_{t+1}, k_{t+1}, I_t) b_{t+1} - d_t - \frac{R}{\alpha} k_{t+1} - I_t - c_f\right].$$

FOCs are given by:

$$[d_t]: \quad (1-\mu_t)\, d_t = 0 \tag{69}$$
$$[k_{t+1}]: \quad \mu_t \left[\frac{R}{\alpha} - \frac{\partial q_t(\cdot)}{\partial k_{t+1}}\, b_{t+1}\right] = \frac{1}{1+r}\mathbb{E}_t\left[\frac{\partial}{\partial k_{t+1}} V_{t+1}^{\text{bank}}(\cdot)\right] \tag{70}$$
$$[I_t]: \quad \mu_t \left[1 - \frac{\partial q_t(\cdot)}{\partial I_t}\, b_{t+1}\right] = \frac{1}{1+r}\mathbb{E}_t\left[\frac{\partial}{\partial I_t} V_{t+1}^{\text{bank}}(\cdot)\right] \tag{71}$$
$$[b_{t+1}]: \quad \mu_t \left[q_t(\cdot) + \frac{\partial q_t(\cdot)}{\partial b_{t+1}}\, b_{t+1}\right] = \frac{-1}{1+r}\mathbb{E}_t\left[\frac{\partial}{\partial b_{t+1}} V_{t+1}^{\text{bank}}(\cdot)\right] \tag{72}$$

By defining $\mu_{t+1}(o_{t+1}, \epsilon_{t+1}, \text{fail}_{t+1}, n_{t+1}; \hat{\theta}) = 0$ if $\Delta_{t+1}(o_{t+1}, \epsilon_{t+1}, \text{fail}_{t+1}, n_{t+1}; \hat{\theta}) = 1$ (i.e., Lagrangian multiplier is zero in a state the firm defaults), I obtain:

$$\frac{\partial}{\partial k_{t+1}} V_{t+1}^{\text{bank}}(\cdot) = \mu_{t+1} \frac{\partial o_{t+1}}{\partial k_{t+1}}$$
$$\frac{\partial}{\partial b_{t+1}} V_{t+1}^{\text{bank}}(\cdot) = -\mu_{t+1}$$

A useful Lemma is shown below:

Lemma 1.
(a) If $P(\Delta_{t+1}) = 0$, $\mu_t = \mathbb{E}_t[\mu_{t+1}] \geq 1$.
(b) If $P(\Delta_{t+1}) > 0$, $\mu_t > 1$.

Proof. (a) If $P(\Delta_{t+1}) = 0$ (zero probability of default in the next period), the price of debt is given by (due to the zero profit condition):

$$q_t(\cdot) = \frac{1}{1+r}.$$

Then, (72) implies

$$\mu_t = \mathbb{E}_t[\mu_{t+1}].$$

When a firm does not default ($\Delta_{t+1} = 0$), the Lagrangian multiplier is greater than 1 ($\mu_{t+1} \geq 1$) due to the Kuhn-Tucker condition (69).

(b) Equation (72) implies that

$$\mu_t = \frac{\frac{1}{1+r}\mathbb{E}_t\left[\mu_{t+1}\right]}{q_t(\cdot) + \frac{\partial q_t(\cdot)}{\partial b_{t+1}}\, b_{t+1}}. \tag{73}$$

By rewriting $q_t(\cdot)$ as:

$$q_t(b_{t+1}, k_{t+1}, I_t) = \frac{1}{1+r}\mathbb{E}_t\left[1 - \mathbf{1}\left\{\Delta_{t+1} = 1\right\}\left(1 - \frac{y_{t+1}}{b_{t+1}} + c_d\right)\right], \tag{74}$$

I obtain:

$$\begin{aligned}\frac{\partial q_t(\cdot)}{\partial b_{t+1}} =& \frac{-1}{1+r}\mathbb{E}_t\left[\mathbf{1}\left\{\Delta_{t+1} = 1\right\}\frac{y_{t+1}}{(b_{t+1})^2}\right] \\ & - \frac{1}{1+r}\mathbb{E}_t\left[\mathbf{1}\left\{\Delta_{t+1} = 1\right\}\left(1 - \frac{y_{t+1}}{b_{t+1}} + c_d\right)\right]\frac{\partial}{\partial b_{t+1}}P(\Delta_{t+1} = 1).\end{aligned}$$

The second term is positive, which will be denoted by $\frac{1}{(1+r)b_{t+1}}\tilde{\Delta} \geq 0$. Equation (73) is given by:

$$\begin{aligned}\mu_t &= \frac{\frac{1}{1+r}\mathbb{E}_t\left[\mu_{t+1}\right]}{\frac{1}{1+r}\mathbb{E}_t\left[1 - \mathbf{1}\left\{\Delta_{t+1} = 1\right\}\left(1 - \frac{y_{t+1}}{b_{t+1}} + c_d\right)\right] - \frac{1}{1+r}\mathbb{E}_t\left[\mathbf{1}\left\{\Delta_{t+1} = 1\right\}\frac{y_{t+1}}{b_{t+1}}\right] - \frac{1}{1+r}\tilde{\Delta}} \\ &= \frac{\mathbb{E}_t\left[\mu_{t+1}\right]}{1 - (1+c_d)P(\Delta_{t+1}) - \tilde{\Delta}}.\end{aligned} \tag{75}$$

Because $\mathbb{E}_t\left[\mu_{t+1}\right]$ is larger than the sum of two cases, $1 \cdot P(\Delta_{t+1} = 0) + 0 \cdot P(\Delta_{t+1} = 1)$,

$$\mu_t \geq \frac{1 - P(\Delta_{t+1} = 1)}{1 - (1+c_d)P(\Delta_{t+1}) - \tilde{\Delta}} > 1.$$

□

From Property (b), if $P(\Delta_{t+1} = 1) > 0$ (or $q_t(\cdot) < \frac{1}{1+r}$), $d_t = 0$ by the Kuhn-Tucker condition (69). Also, $\mu_t = 1$ implies $P(\Delta_{t+1} = 1) = 0$ and $q_t(\cdot) = \frac{1}{1+r}$.[82]

Value Function Iteration The value function is computed separately at the last step of the productivity ladder ($n = N$), where firms do not invest in innovation, and at $n < N$, where firms invest in innovation. In addition, the firm's decision is separately analyzed for two cases: $d_t > 0$ and $d_t = 0$.

If $d_t > 0$ (i.e., firms distribute dividends), the Lagrangian multiplier is one ($\mu_t = 1$) implied by the Kuhn-Tucker condition (69). If $d_t = 0$, the Lagrangian multiplier (the

---

[82] If $P(\Delta_{t+1} = 1) = 0$ and $q_t(\cdot) = \frac{1}{1+r}$, firms could borrow money from banks at a risk-free rate and distribute it as dividend, which does not change the firm value. Such solutions are eliminated by assumption in the value function iteration. On the other hand, if firms may face a bank loan rate higher than a risk-free rate in the next period, firms have an incentive to save this period. Firms would not choose $d_t > 0$ if $\mathbb{E}\left[\mu_{t+1}\right] > 1$, since firms increase savings ($b_t < 0$) until $d_t = 0$ or $\mathbb{E}[\mu_{t+1}] = 1$ is satisfied. In the firm dynamics literature, the issue of indeterminacy of debt is addressed, for instance, by introducing tax incentives for debt (Corbae and D'Erasmo, 2021), impatient households (Cooley and Quadrini, 2001), or positive default probability on all occasions (Khan et al., 2014).

shadow price of cash flow) is larger than or equal to one. Note that the pricing function $q_t(\cdot)$ is computed given the value function in the next period:

$$\begin{aligned}
&q_t(b_{t+1}, k_{t+1}, I_t; n_t, \epsilon_t, \text{fail}_t, V_{t+1}(\cdot)) \\
&= \frac{1}{1+r}\left[P(\Delta_{t+1}=0) + P(\Delta_{t+1}=1)\frac{\mathbb{E}_t\left[y_{t+1} - c_d b_{t+1} | \Delta_{t+1}=1\right]}{b_{t+1}}\right] \\
&= \frac{1}{1+r}\mathbb{E}_t\Bigg[(1-\varphi)\big(\mathbf{1}\left\{V_{t+1}(\cdot) \geq 0\right\} + \mathbf{1}\left\{V_{t+1}(\cdot) < 0 \wedge o_{t+1} \geq 0\right\} \\
&\qquad + \mathbf{1}\left\{V_{t+1}(\cdot) < 0 \wedge o_{t+1} < 0\right\}\left(\frac{y_{t+1}}{b_{t+1}} - c_d\right)\big) \\
&\qquad + \varphi\left(\mathbf{1}\left\{o_{t+1} \geq 0\right\} + \mathbf{1}\left\{o_{t+1} < 0\right\}\left(\frac{y_{t+1}}{b_{t+1}} - c_d\right)\right)\Bigg]. \qquad (76)
\end{aligned}$$

The budget constraint is given by:

$$d_t + \frac{R}{\alpha}k_{t+1} + I_t + c_f = o_t + q_t(b_{t+1}, k_{t+1}, I_t)b_{t+1}. \qquad (77)$$

I. Consider the case at $n = N$. In this case, since firms are in the last step of the productivity ladder, they do not invest in innovation, and a set of state variables is given by (o_t, ϵ_t). In each iteration, the optimal choice of (k_t, d_t, b_{t+1}) and the value function $V_t(\cdot)$ are computed, given the value function in the next period $V_{t+1}(\cdot)$, Lagrangian multiplier $\mu_t(\cdot)$, and pricing function $q_t(\cdot)$.

I–(i). A case with $d_t > 0$ and an unconstrained solution k_t^*:
If a firm has large enough cash flow o_t with Lagrangian multiplier $\mu_t = 1$, the unconstrained solution of capital k_t^* is chosen:

$$k_t^* = \mathbb{E}_t\left[z_{t+1}^{1-\eta}\right]^{\frac{1}{1-\eta}}\left\{\frac{\eta}{1+r}\left(\frac{\alpha}{R}\right)^{1-(1-\alpha)\eta}\left(\frac{1-\alpha}{w}\right)^{(1-\alpha)\eta}\right\}^{\frac{1}{1-\eta}}. \qquad (78)$$

For each value of cash flow (o_t), the unconstrained solution of capital satisfies the budget constraint if:

$$o_t > \left(\frac{R}{\alpha}\right)k_t^* + c_f. \qquad (79)$$

If this is the case, the solution is given by:

$$\begin{aligned}
k_t &= k_t^* \\
d_t &= o_t - \left(\frac{R}{\alpha}\right)k_t^* - c_f \\
b_t &= 0 \\
V_t(\cdot) &= d_t + (1-\varphi)\frac{1}{1+r}\mathbb{E}\left[V_{t+1}(o_{t+1}, \epsilon_{t+1}, \text{fail}_{t+1}=0, n_{t+1}=N)\right] + \varphi\frac{1}{1+r}\mathbb{E}\left[o_{t+1}\right]
\end{aligned}$$

I–(ii). A case with $d_t = 0$ and $b_t > 0$:
If the inequality (79) does not hold, a firm cannot purchase sufficient capital without

raising a bank loan. The optimal amount of capital is chosen subject to the budget constraint (77). The dept pricing function $q_t(\cdot)$ is determined by equation (76), and the size of the debt and capital expenditures are determined to satisfy the budget constraint. Numerically, the firm value is computed with several values of k_{t+1} from grid points, and the optimal k_{t+1} (and corresponding b_{t+1} determined by the budget constraint) that maximizes the firm value is chosen. Given the optimal (k_{t+1}, b_{t+1}) and $d_t = 0$, the firm value is given by:

$$V_t(\cdot) = (1-\varphi)\frac{1}{1+r}\mathbb{E}\left[V_{t+1}(o_{t+1}, \epsilon_{t+1}, \text{fail}_{t+1} = 0, n_{t+1} = N)\right] + \varphi\frac{1}{1+r}\mathbb{E}\left[o_{t+1}\right]$$
$$\text{with } o_{t+1} = y_{t+1} - b_{t+1}.$$

If the firm value is negative with any choice of k_{t+1} and cash flow is negative ($o_t < 0$), the firm defaults with the outstanding bank loan and exits the market. If cash flow is positive, $o_t > 0$, but the firm value is less than the cash flow with any choice capital, the firm exits the market after fully repaying the loan.

The value function is iterated, meaning that the value function $V_t(\cdot)$ is computed given an updated value function in the next period $V_{t+1}(\cdot)$, until it converges.[83]

II. Consider the case at $n < N$. In this case, firms invest in innovation.

II-(i). A case with $d_t > 0$ and an unconstrained solution (k_t^*, I_t^*):

The unconstrained solution (k_t^*, I_t^*) follows a system of two equations. The first equation is the same as equation (78), where the realization of productivity z_{t+1} involves uncertainty on innovation outcome and an exogenous productivity shock:

$$\mathbb{E}_t\left[z_{t+1}^{1-\eta}\right] = \mathbb{E}_t\left[\sigma_{t+1}(\theta_{n_t+1} + \epsilon_{t+1})^{1-\eta} + (1-\sigma_{t+1})(\theta_{n_t} + \epsilon_{t+1})^{1-\eta}\right]. \tag{80}$$

The FOC with respect to innovation investment (I_t) is given in equation (71):

$$\mu_t\left[1 - \frac{\partial q_t(\cdot)}{\partial I_t}\, b_{t+1}\right] = \frac{1}{1+r}\mathbb{E}_t\left[\frac{\partial}{\partial I_t}V_{t+1}^{\text{bank}}(\cdot)\right]$$

The derivative of the innovation function gives:

$$\frac{\partial \varsigma}{\partial I} = \varsigma(1-\varsigma)\frac{\gamma_1}{I}. \tag{81}$$

Using this, the two terms inside the equation (71) are derived as:[84]

$$\mathbb{E}_t\left[\frac{\partial}{\partial I_t}V_{t+1}^{\text{bank}}(\cdot)\right] = \varsigma^*(1-\varsigma^*)\frac{\gamma_1}{I_t}\mathbb{E}_t\left[V_{t+1}^{\text{bank}}(o_{t+1}, \epsilon_{t+1}, \text{suc}, n_t+1) - V_{t+1}^{\text{bank}}(o_{t+1}, \epsilon_{t+1}, \text{fail}, n_t)\right]$$
$$\frac{\partial q_t(\cdot)}{\partial I_t} = \varsigma^*(1-\varsigma^*)\frac{\gamma_1}{I_t}\frac{-1}{1+r}\mathbb{E}_t\left[\Delta_{t+1(n_t+1)}\left(1 - \frac{y_{t+1}(n_t+1)}{b_{t+1}} + c_d\right) - \Delta_{t+1(n_t)}\left(1 - \frac{y_{t+1}(n_t)}{b_{t+1}} + c_d\right)\right]$$

---

[83] The Lagrangian multiplier and pricing function are updated once in several iterations for the value function iteration to be stable. The value function is iterated several times with the same choice of control variables before reoptimizing the control variables to achieve faster convergence.

[84] $q_t(\cdot)$ is given by (74): $q_t(b_{t+1}, k_{t+1}, I_t) = \frac{1}{1+r}\mathbb{E}_t\left[1 - \mathbf{1}\{\Delta_{t+1} = 1\}\left(1 - \frac{y_{t+1}}{b_{t+1}} + c_d\right)\right]$

The unconstrained solution (k_t^*, I_t^*) is obtained by solving a system of equations (78) and (71) with $\mu_t = 1$. If it satisfies the budget constraint given the firm's cash flow:

$$o_t > \left(\frac{R}{\alpha}\right) k_t^* + I_t^* + c_f, \tag{82}$$

the solution is given by:

$$\begin{aligned}
k_t &= k_t^* \\
I_t &= I_t^* \\
d_t &= o_t - \left(\frac{R}{\alpha}\right) k_t^* - I_t^* - c_f \\
b_t &= 0 \\
V_t(\cdot) &= d_t + (1-\varphi)\frac{1}{1+r}\mathbb{E}\left[V_{t+1}(o_{t+1}, \epsilon_{t+1}, \text{fail}_{t+1}, n_{t+1})\right] + \varphi\frac{1}{1+r}\mathbb{E}\left[o_{t+1}\right].
\end{aligned}$$

II-(ii). A case with $d_t = 0$ and $b_t > 0$:
If the unconstrained solution is not feasible with the given cash flow o_t, firms raise a bank loan and optimally choose capital and innovation expenditure subject to the budget constraint (77). The firm value is evaluated at several values of k_t from grid points, and the optimal k_t that maximizes the firm value is chosen. For each k_t, the optimal (I_t, b_{t+1}) is solved to satisfy the FOC (71) and the budget constraint (77). The system of two equations is solved with NLsolve and Optim commands in Julia.[85] Given the optimal choice of (k_t, I_t, b_{t+1}), the value function is updated and iterated until it converges. The value function is iterated several times with the same choice of control variables before reoptimizing the control variables.

B.3 Additional Discussions

B.3.1 Additional Quantitative Results

Figure 16 shows the relation between a firm's growth potential (θ) and its bank default rate at age 2 (one period after its birth). Although high-growth and medium-potential firms optimally choose VC financing and Angel financing in equilibrium, the figure shows the hypothetical bankruptcy rate if all firms choose bank financing. High-potential and medium-potential firms will default at a higher rate than lower-potential firms because higher-potential firms borrow more to finance innovation, but the innovation outcome is uncertain. Because of the high default rate, higher-potential firms face a higher bank loan rate, as seen in the higher credit spread (i.e., bank default rate minus risk-free rate) in the right panel.

---

[85] Several efforts are made to speed up computation: (i) the solution from the last iteration is used as an initial value, (ii) a value of k' in the grid is eliminated from the potential solution if it cannot satisfy the budget constraint with the highest value of $q(\cdot)b'$, (iii) once the value function passes the modal point with respect to k' (firm value first increases with k' and declines), a firm value at higher k' is not evaluated.

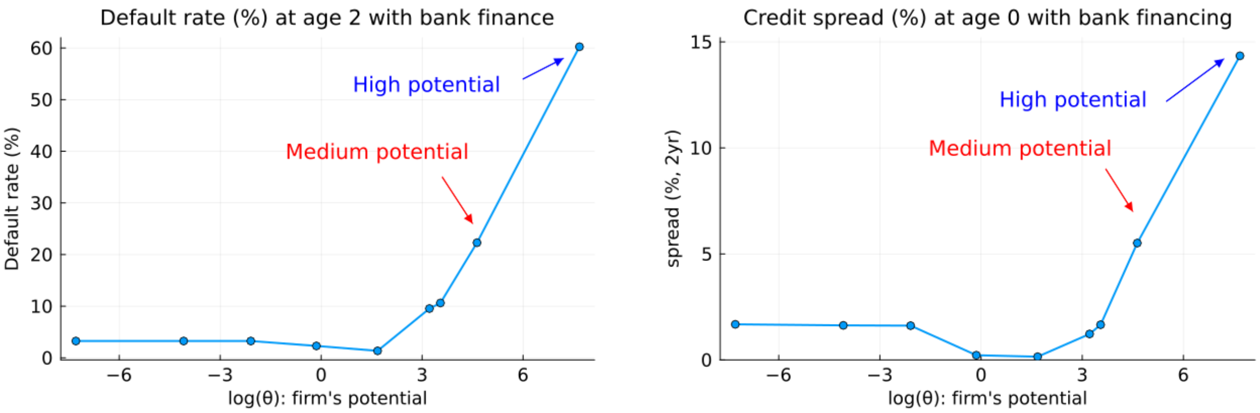


Figure 16: Default rate and Credit Spread for each θ

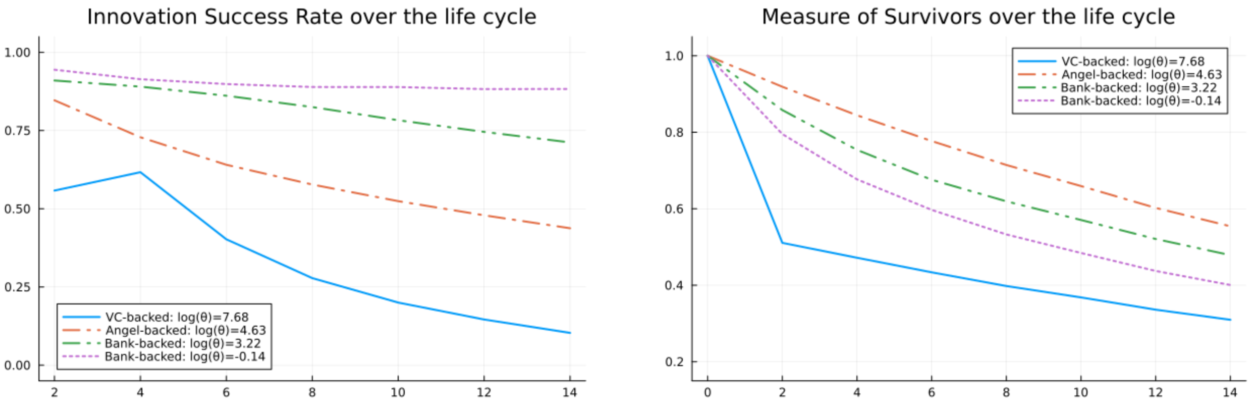


Figure 17: Lifecycle Dynamics

Figure 17 shows the lifecycle dynamics of innovation success rate and the measure of survivors for each type of firm. The calibrated model implies that higher-potential firms have a lower innovation success rate than lower-potential firms. In the right panel, around half of VC-backed firms exit at age 2. This is because the firm value of VC-backed firms that failed in innovation is too low, given the management fees paid to the venture capitalists. Among the bank-backed firms, higher-potential firms are more likely to survive, given their higher firm values.

B.3.2 Note on the Tauchen Method

An exogenous productivity shock $\log(\epsilon_t)$ follows an AR(1) process. In the VFI, the stochastic process is approximated by the Tauchen's method. To reduce the computation time, a small number of grid points for $\log(\epsilon_t)$ is chosen (e.g. $N_\epsilon = 5$). In the standard Tauchen's method with a small number of grid points, the probability distribution across grid points is highly unequal, and the values of ϵ may not change frequently across periods. This could lead to an unreasonably low exit rate in the model. To address this issue, the grid points are chosen so that the probability of each grid point in the stationary distribution is approximately equal.

To be specific, the grid points of $\log(\bar{\epsilon}_k)$ for $k = 1, \cdots, N_\epsilon$ are set such that the following relationship holds:

$$P\left(\frac{\sigma_\nu}{\sqrt{1-\rho^2}}\,\nu \le \log(\bar{\epsilon}_i) + \frac{d_i}{2}\right) = \frac{i}{N_k} \text{ for } i = 1, \cdots, N_\epsilon - 1, \tag{83}$$
$$\text{where } d_i = \log(\bar{\epsilon}_{i+1}) - \log(\bar{\epsilon}_i).$$

This is because the variance of $\log(\epsilon_t)$ in the stationary distribution is given by $\left(\frac{\sigma_\nu}{\sqrt{1-\rho^2}}\right)^2$. Thus, the cumulative probability at each grid point is approximately equal to $\frac{i}{N_k}$ for $i = 1, \cdots, N_\epsilon - 1$. The number of grid points, N_k, is chosen to be an odd number, and the grid is symmetric around the middle point, $\log\left(\epsilon_{\frac{1}{2}(N_k+1)}\right) = 0$. Once the grid points are chosen, the transition probability is computed as:

$$P(\epsilon_t = \bar{\epsilon}_k | \epsilon_{t-1} = \bar{\epsilon}_j) = P\left(\log(\bar{\epsilon}_k) - \frac{d_{k-1}}{2} \le \rho\log(\bar{\epsilon}_j) + \sigma_\nu\nu \le \log(\bar{\epsilon}_k) + \frac{d_k}{2}\right) \text{ for } k = 2, \cdots, N_\epsilon - 1,$$

while the transition probability at the first and last grid point is computed as:

$$P(\epsilon_t = \bar{\epsilon}_1 | \epsilon_{t-1} = \bar{\epsilon}_j) = P\left(\rho\bar{\epsilon}_j + \sigma_\nu\nu \le \bar{\epsilon}_1 + \frac{d_1}{2}\right), \tag{84}$$

$$P(\epsilon_t = \bar{\epsilon}_{N_\epsilon} | \epsilon_{t-1} = \bar{\epsilon}_j) = P\left(\bar{\epsilon}_{N_\epsilon} - \frac{d_{N_\epsilon-1}}{2} \le \rho\bar{\epsilon}_j + \sigma_\nu\nu\right). \tag{85}$$

Using this method, the probability at ϵ_3 in case of $N_\epsilon = 5$ is 0.22, while in the standard method, the probability at ϵ_3 is 0.37 if a multiple of two of the unconditional standard deviation is chosen for ϵ_{ϵ_N}.

B.3.3 A note on stationary distribution

A stationary distribution is computed by simulation. The stationary distribution Ψ^* is defined over state variables $h := (n, \epsilon, o, \hat{\theta})$.

Given policy functions obtained from the value function iteration, it is straightforward to simulate an evolution of state variables over the firm's life cycle. This is done by drawing the initial state $h_0 = (n_0, \epsilon_0, o_{\text{initial}}, \hat{\theta})$ and realization of shocks (ϵ_t and innovation outcome) over the life cycle. The control variables $(\ell_t, k_t, I_t, b_t, d_t)$ are determined by the policy functions conditional on the firm's state variables.

Although firm age is not a part of state variables, the stationary distribution can be expressed as a sum of distributions at each age:

$$\Psi^*(h) = \sum_{a=0}^{\bar{a}} \Psi^*(h; a), \tag{86}$$

where a denotes firm age from 0 to $\bar{a}$, and $\bar{a}$ is the maximum firm age. A large enough value for $\bar{a}$ is chosen. Note that a stationary distribution is typically computed by drawing an arbitrary initial distribution Ψ^0 and sequentially updating the distribution until it converges based on a transition function P:

$$\Psi^{j+1} = P(\Psi^j). \tag{87}$$

Here, using the distribution conditional on age, the distribution at iteration j is expressed as:

$$\begin{aligned} \Psi^j &:= \sum_{a=0}^{\bar{a}} \Psi^j(h; a) \\ &= m\bar{\psi} + \sum_{a=1}^{\bar{a}} P_{a-1\to a}\left(\Psi^{j-1}(h; a-1)\right), \end{aligned} \tag{88}$$

where $\bar{\psi}$ denotes the distibution of entrants. A function $\Psi^{j-1}(\cdot; a-1)$ computes a distribution $\Psi^j(\cdot; a)$ of iteration j at firm age a from a distirbution $\Psi^{j-1}(\cdot; a-1)$ at firm age $a-1$. Notice that the distribution Ψ^j at age a can be computed by a simulation of many firms based on the distribution at age $a-1$. Some firms exit the market, and all of their control variables are set to zero from the period onwards. The measure of the entrants is given by m and is a constant in the steady state. By taking a large j that satisfies $j > \bar{a}$, Ψ^j is expressed as:

$$\Psi^j = m\bar{\psi} + P_{0\to 1}\left(m\bar{\psi}\right) + P_{0\to 2}\left(m\bar{\psi}\right) + \cdots + P_{0\to \bar{a}}\left(m\bar{\psi}\right). \tag{89}$$

This implies that the stationary distribution can be derived as a sum of distributions across all ages from $a = 0$ to $a = \bar{a}$, given the initial distribution of entrants ψ. The distribution of each age group $P_{0\to a}\left(m\bar{\psi}\right)$ can be simulated given the initial distribution of entrants.