EFRSW
7/18/06
3:30 PM

On the Return to Venture Capital*

Boyan Jovanovic[†] and Balázs Szentes[‡]

July 11, 2006

Abstract

We argue that the seemingly sizeable excess return to venture equity arises because of a relative shortage of venture capitalists – people who have the expertise to assess the profitability of projects, and have liquidity to finance them. A VC must supply not only money but also his time and it is the scarcity of the latter that keeps the return to venture equity high. The VC is therefore less patient with maturing firms than an ordinary entrepreneur would be, and this may explain why venture-backed reach IPOs earlier than other start-ups and why they are worth more at IPO.

The scarcity of VCs enables them to internalize their social value, so that the competitive equilibrium is socially optimal. This optimality obtains on an open set of parameter values.

We estimate the model and back out the return of solo entrepreneurs which is always below that of the return on venture equity.

1 Introduction

Venture Equity (VE) funds earn returns of several points in excess of what their risk characteristics would warrant. Surveying recent results, Kaplan and Schoar (2005, henceforth KS) report that VE-funds' alphas tend to be around four or five percent, though the estimates vary from -1.5% to 32%.[1] The standard errors around these

*We thank I. Guler, and Y. Yafeh for comments and for providing us with data, M.A. Campo-Rembado, W. Fuchs, C. Michelaci, M. Ueda and, especially, A. Ljungqvist for comments, the NSF for support and A. Gavazza, V. Tsyrennikov, and M. Kredler for ably estimating the model and providing comments on the entire paper.

†Department of Economics, NYU. Address: 269 Mercer Street, New York, NY 10003. E-mail: Boyan.Jovanovic@nyu.edu.

‡Department of Economics, University of Chicago. Address: 1126 59th Street, Chicago, IL 60657. E-mail: szentes@uchicago.edu.

[1] The studies they discuss and the alphas that they report are Cochrane (2005) 0.32% Jones and Rhodes-Kropf, (2004) 4.7%, and Ljungqvist and Richardson (2003) $\alpha > 0$. Experienced VCs gnerate higher returns – KS as well as Lerner, Schoar and Wong (2005) find this.

estimates are large, moreover, as the aggregate-risk component of VE funds is hard to determine because of their illiquidity. Overall, however, we may take it as fairly well established that the rates of return to VE are above normal, and that their alphas are perhaps in the range of four to five percent.

If VE returns are above normal, why isn't there more such equity supplied until the alphas are driven to zero through some sort of diminishing returns? One answer is that it requires a 5 to10-year lock-up of funds that a small investor cannot commit to. VE should be attractive to institutional investors, however, since they are liquid and well diversified. Rather, the answer seems to be that the number of VE funds is limited not by a scarcity of investors or their capital, but by a scarcity of VCs.

We model a market for venture capital in which

1. VC scarcity raises the effective discount rate of the VC. The higher the excess return to venture capital, the more impatient a VC is with the companies in his portfolio and the earlier he terminates a non-profitable company.[2]

2. The outcome is nonetheless efficient for a wide range of parameter values. Efficiency pertains to both (*i*) Contracting between VCs and entrepreneurs and the resulting termination rules, and to (*ii*) Entrepreneurs' choices over whether to seek venture backing, whether to develop the project on their own, or whether to abandon the project altogether.

Sketch of the model.—The main actors in the model are VCs who are endowed with unlimited wealth, and entrepreneurs who are endowed with limited wealth and with projects. A project entails start-up costs and continuation costs. Whether it is backed by a VC or not, the start-up cost must be paid before any information about its quality can come in. After that, continuation costs must be paid until the project succeeds or is terminated. Start-up costs entail only capital, but continuation costs entail capital and effort: funds must be supplied and the entrepreneur must exert effort without interruption until the project yields fruit. Project quality in the model has two dimensions: The size of the return, and the waiting time until the return is realized. Neither dimension is known before a contract between a VC and an entrepreneur is signed. After the contract is signed, however, some uncertainty is resolved. After that, either party can, at any time, terminate the project. The entrepreneur can do so by withholding effort, and the VC can do so by withholding capital. The optimal contract is set up so that when a project is terminated, both are better off: The entrepreneur no longer wishes to exert effort, and the VC no longer wishes to lend. If the entrepreneur develops the project on her own, she may run out

[2] Jones and Rhodes-Kropf (2004) argue that VCs use discount rates as high as 30 to 50 percent. A different hypothesis from the behavioral finance literature, is that the VC can neutralize the company's founder's irrational attachment to the project at hand, that he protects the limited partners in the private equity fund from losses on non-performing portfolio companies.

of money before the optimal termination date. A poor entrepreneur therefore must either use venture funding (as bank lending is infeasible because the bank cannot recognize some critical signals about the quality of the project) or give up on the project.

The efficiency result.—We choose parameter values so that VCs' time endowment is smaller than the number of projects that demand VC services. Since VCs reject most of the proposals that they receive, this assumption seems reasonable. As a result, all the rents from the VC-entrepreneur partnership go to the VC, and this allows VCs to internalize their social value. In this way we get the equilibrium to be efficient.

Notes on the literature.—Bergemann and Hege (1998, forthcoming), deal with a single VC and a single entrepreneur, with their outside options taken as given. Holmes and Schmitz (1990) analyze a market equilibrium and determine the rewards of founders relative to managers of firms, but do not analyze venture capital. Muller and Inderst (2004) model the market for venture capital as do Michelacci and Suarez (2004) who, in addition, analyze the termination decision and link it to the equilibrium value of a VC. Later, we shall discuss our results in relation to these two papers – neither paper analyzes the entrepreneur's decision of whether to seek venture backing or to finance the project in some other way. Ueda (2004) analyze this mode-of-financing decision in a model where one cost of VC financing is the threat that the VC use the information to set up a competing business. We take the wealth of an entrepreneur as given, whereas Buera (2004) and Basaluzzo (2004) analyze the saving behavior of future entrepreneurs. Cochrane (2005) deals with the pricing the income streams to VC portfolios; unlike us, he allows for risk aversion, but he does not derive optimal termination rules. In our model, a VC has expertise and an infinitely deep pocket, so that he does not need to seek outside funds. We therefore do not address the question of how rents are divided between VCs and outside investors and, hence, we do not solve the puzzle of why the rents do not *all* go to VCs.[3] Jones and Rhodes-Kropf (2004) focus on this aspect, emphasizing the agency problem between VCs and outside investors. They argue that, in order to induce him to exert effort, the VE investors must compensate the VC via an equity interest, which forces the VC to hold a lot of idiosyncratic risk.

Plan of the paper.—The next section describes the model, and Section 3 derives the equilibrium contract and shows that the competitive outcome is efficient. Section 4 derives several empirical implications of the model and discusses evidence. Section 5 solves an example by hand and fits it to longitudinal data on VC investments, spanning 1989-2000, and their performance outcomes. Section 6 concludes the paper

[3] KS point out that the VCs do not appropriate the returns to their skill via higher fees and larger funds. Gompers and Lerner (1999) found that the VCs that performed well also tended to have higher profit shares, and KS find that the best VC funds tended to raise their profit shares more recently. Nevertheless KS find these trends have not been strong enough to eliminated the persistence in fund returns, i.e., too weak to eliminate the VE premium.

and the Appendix describes the data and the estimation procedure.

2 Model

There is a measure x of infinitely lived VCs, each able to borrow unlimited amounts of money at the rate r. There if also an inflow at the rate λ of potential projects, each in the possession of a different entrepreneur. The entrepreneurs cannot borrow, and have initial wealth w which is distributed according to the CDF $\Psi(w)$. An entrepreneur can have at most one idea, ever.

A Project

A project can be undertaken by an entrepreneur alone, in which case she must rely on her own wealth only, or together with a VC. For the project to succeed, it requires an immediate payment of a cost C, and after that it also requires k units of investment and a units of effort by the entrepreneur at every instant up until the project yields a return. The project yields a return π at time τ, where both π and τ are random variables, independent of one another.[4] Let F denote the distribution of τ, and f the corresponding density. Let h denote the hazard rate corresponding to F, that is $h = f/(1-F)$. We assume that the hazard rate h has a bell-shape. It first increases, then decreases. In other words, as time passes without the realization of π, the agents first become more optimistic about a quick realization of π, but then they become more and more pessimistic.[5] These assumptions seem to fit the facts at least roughly; Lerner (1998, p. 738) writes:

> Immediately after a new venture is financed, the probability that there will be significant information inflows is actually likely to be quite low: the entrepreneur is in all probability focusing on the early development of his businesses. At some point thereafter, however, the probability that information will arrive increases dramatically: e.g., the results of the clinical trial will emerge, the prototype will be either be successfully developed or not, or the manufacturing yields from the new production line will become known.

If the project is either not invested into or effort is not exerted, the project cannot yield a positive return, ever. Neither party knows π and τ, but their distributions are common knowledge. In a venture-backed firm, after the contract is signed and after a cost C is incurred, π becomes known to both parties. This is where the VC has the advantage over a bank which lacks the necessary expertise and cannot learn π before date τ. However, no information about τ is received. In a solo venture, the

[4] The independence assumption simplifies the algebra but is inessential for the results until we reach the prediction summarized in Figure 6. We shall coment further on it then.

[5] Our theoretical results also hold if the hazard declines monotonically throughout.

entrepreneur alone incurs C at the outset, and thereby she learns π. Since the solo entrepreneur also has to pay C, it is not a project-screening cost but should instead be thought of as a lumpy initial investment.

Let G denote the distribution of π, and g the corresponding density. The expected social value to developing projects is assumed to be positive, once the option to stop the project at some point is taken into account.

Preferences

The entrepreneur and the VC are risk neutral and both discount the future at the rate r. The VC maximizes the expected discounted present value of his net income. The entrepreneur maximizes the expected discounted present value of her income minus her disutility, a_t, from exerting effort.

We choose units of a and k so that the amounts required to keep the project alive sum up to one: $a + k = 1$. This normalization has no bearing on the analysis because a doubling of all costs and benefits leaves unchanged all the variables that we shall consider, namely the duration of projects and their rates of return.

Market Structure

When an entrepreneur gets an idea, she has to decide whether to abandon her project, to seek VC-backing, or to go solo, i.e., to implement her project alone. This decision is irreversible.

Suppose at time t there is a measure n of VCs who is not in a contractual relationship with entrepreneurs and a measure of m of entrepreneurs who wishes to be financed by a VC. Then the number $\min\{n, m\}$ of VCs and entrepreneurs are randomly matched and can enter into a contractual relationship.

Timing

Events occur in the following sequence:

1. Entrepreneur chooses whether to (i) invest her wealth with a bank, (ii) develop her project on her own, or (iii) sign a contract with a VC,

2. Under option (ii) or (iii), pay a cost C,

3. π is then revealed,[6]

4. No signals come in about τ until it is realized.

Contracting

Feasible Contracts —If the entrepreneur chooses (iii), the contract the VC can offer consists of two positive numbers: (p, s). The number p is an up-front payment

[6] There is no prior signal about π.

the VC pays the entrepreneur right after signing a contract. This means that the entrepreneur in that case bears only $C - p$ of the up-front cost. The number s, specifies how to share the return if the project succeeds. If the project yields return π, the entrepreneur gets $s\pi$ and the VC gets $(1-s)\,\pi$. Neither the effort of the entrepreneur nor the subsequent investment of the VC can be contracted on. On the other hand the payments p and $s\pi$ are enforceable.

After the transfer p, this is a pure equity contract. We could allow for more complicated contracts in which s would depend on τ and π. We shall show, however, that the simpler contracts already induce socially efficient decisions. Moreover, the equilibrium outcome of a game with more complicated contracts would be identical to ours.

Timing of the Contractual Relationship —First, the VC offers a contract, (p, s), to the entrepreneur. If the entrepreneur refuses the contract the game between these two parties ends; the entrepreneur has to leave the market and invest with a bank, the VC seeks to be matched with an other entrepreneur. If the entrepreneur signs the contract she receives p from the VC up front. We interpret p as the amount that the VC pays towards financing C. The entrepreneur finances the remaining part of C and both parties then immediately learn the value of π.

After a length of time t, if the return has not yet been realized, both parties must decide whether to continue supporting the project or not. That is, the entrepreneur has to decide whether to exert effort and the VC has to decide whether to invest. One can assume that the parties can observe the history of investments and effort up to time t, when making these decisions.[7] If either party decides not to support the project the game between the two parties ends, otherwise it continues. When it does end, the VC is free to devote his time to another project. The entrepreneur, on the other hand, must leave the market – there are no serial entrepreneurs in the model.

If the project yields a return π at time t, the entrepreneur gets $s\pi$ the VC gets $(1-s)\,\pi$ and the game ends between the two parties. Again, the VC seeks to be matched with new entrepreneurs, and the entrepreneur leaves the market. There are, in other words, no "serial entrepreneurs", only serial VCs.

Banks

In our model, the only role of the banks is to guarantee a risk-free interest rate, but they do not finance projects. This is because VCs are assumed to have two advantages over banks. First, banks lack the expertise of the VCs which is necessary to learn π after paying the cost C. Hence banks can only learn π at the date of success, τ, but not before. Second, banks also lack the monitoring ability of the VCs which ensures that the entrepreneurs do not divert investment to private consumption. As a result, banks do not offer contracts to entrepreneurs, for otherwise anybody could pretend

[7] So as to avoid coordination problems, we assume that at time t the VC observes the history of efforts on the interval $[0, t]$ and that the entrepreneur observes the history of investments on the interval $[0, t)$.

to be an entrepreneur, divert the borrowed funds for personal use, and banks would make negative profits.

Comment on the informational structure.—The assumption that no one can learn about π or τ until someone has financed C reflects the reality that many projects never reach IPO or acquisition and entail large losses. On its own, this requires that either π or τ, or perhaps both, are not known before costs are sunk. Since many projects are terminated in year 5, 6, and even beyond, it is clear that VCs do not know τ and that they give up hope at some point. The later terminations are presumably for projects that promise a large π in the event of success – the model will explain the dispersion in termination dates with a dispersion in (perfect) signals about π. The fairly large spike in terminations during the first year suggests that fairly good signals about π are available fairly early so that the bad apples can be detected and thrown out right away.

3 Analysis

First, we characterize the socially optimal outcome of our model. Then we show that this outcome is the unique outcome in the competitive market conditional on some distributional assumption on the wealth of the entrepreneurs.

3.1 Socially Optimal Decisions

We proceed as follows. First, we analyze the optimal decision regarding how long a project should be supported. This decision depends on whether the project is venture-backed , or supported by a solo entrepreneur. Second, we characterize the socially optimal decision whether an entrepreneur should go solo, seek VC-backing, or invest with a bank.

3.1.1 The termination problem of a venture-backed project

The VC has "unlimited wealth" the transfer of which over periods he values at the market rate of interest r. He can, however, handle only one project at a time. How long should the VC and the entrepreneur support a project? Since the value π is revealed at time zero, we derive the optimal time until the project should be supported, denoted by $T^*(\pi)$.

Let W denote the social value of a free VC. Once π is known and C has been sunk, the planner solves

$$V(\pi) \equiv \max_T \int_0^T \left(\pi + W - \frac{1}{h(t)}\right) e^{-rt} f(t)\, dt + e^{-rT}(1 - F[T])\, W \tag{1}$$

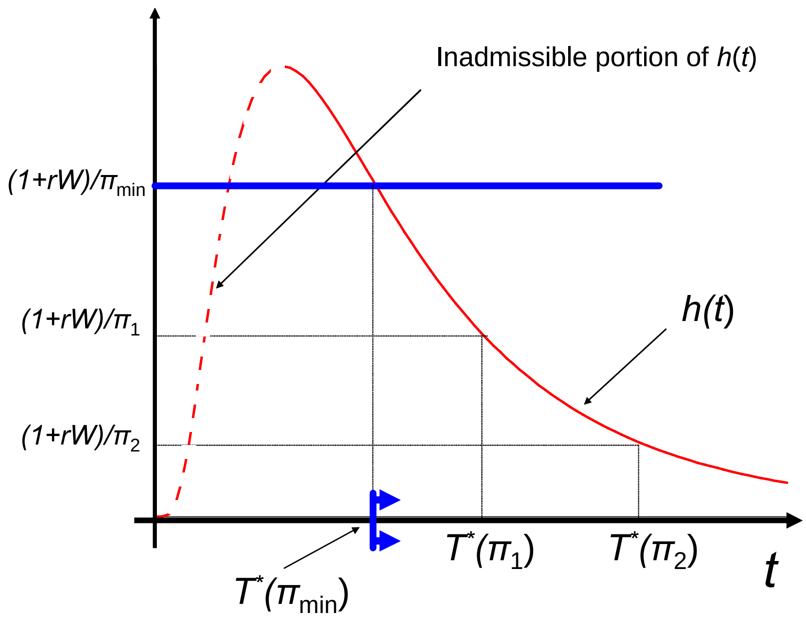


Figure 1: THE DETERMINATION OF $T^*(\pi)$

An interior solution for this problem, $T^*(\pi)$, must solve for t the first-order condition

$$0 = \left(\pi + W - \frac{1}{h(t)}\right) f(t) - f(t) W - r(1 - F[t]) W = \pi - \frac{1 + rW}{h(t)},$$

i.e.,

$$h(t) = \frac{1 + rW}{\pi}.$$

The local second-order condition is $h'(T^*(\pi)) < 0$. The bell-shaped hazard rate assumption guarantees that the local second-order condition is also sufficient, as shown in Figure 1.

Immediate terminations.—Notice however, that T may be at a corner: π may be so low that the project yields a negative return. Let $\pi_{\min}$ be the smallest value of π for which it is worth supporting the project. A project should be funded if and only if $V(\pi) \geq W$. That is, $\pi_{\min}$ solves

$$V(\pi) = W. \tag{2}$$

Therefore the optimal termination age, $T^*(\pi)$, for a project of quality π, is defined as

$$T^*(\pi) = \begin{cases} h^{-1}\left(\frac{1+rW}{\pi}\right) & \text{if } \pi > \pi_{\min}, \\ 0 & \text{otherwise.} \end{cases} \tag{3}$$

3.1.2 The solo entrepreneur's termination problem

For the solo project we assume that π is also drawn from the same G as the venture-backed projects' π, and that it is drawn independently of w; in doing so we implicitly shut off any influence that the entrepreneur's wealth may exert on the scale of businesses.

The unconstrained entrepreneur.—First we solve for the optimal stopping time of a solo entrepreneur, $T^S(\pi)$, who has enough money to finance her project forever. Then $T^S(\pi)$ solves the following maximization problem:

$$\max_T \int_0^T \left(\pi - \frac{1}{h(t)}\right) e^{-rt} f(t)\, dt.$$

Hence, $T^S(\pi)$ is either equal to zero (if the value of the previous maximization problem is negative), or satisfies the first-order condition, $h\left(T^S(\pi)\right) = 1/\pi$, along with the second-order condition $h'\left(T^S[\pi]\right) < 0$. Let $\pi^S_{\min}$ denote the smallest realization of π, which should be supported by a solo entrepreneur with no budget constraint. Hence,

$$T^S(\pi) = \begin{cases} h^{-1}\left(\frac{1}{\pi}\right) & \text{if } \pi > \pi^S_{\min}, \\ 0 & \text{otherwise.} \end{cases} \tag{4}$$

The wealth-constrained entrepreneur.—Given the linear preferences over consumption, the entrepreneur would defer all her consumption until the project is completed. Her wealth then follows the law of motion

$$\frac{dw_t}{dt} = rw_t - k.$$

for $t < T$, where w_t denotes the wealth of the entrepreneur's wealth at time t. The initial condition for wealth is $w_0 - C$. The solution for w_t is

$$w_t = \frac{k}{r} + \left(w_0 - C - \frac{k}{r}\right) e^{rt}. \tag{5}$$

Let $\tau(w)$ be the date at which a solo entrepreneur's wealth runs out conditional on no success until then; $\tau(w)$ is the project's maximum financial life if it does not succeed. Then $\tau(w)$ solves for t the equation $w_t = 0$. Its solution is

$$\tau(w) = \begin{cases} \frac{1}{r}\ln\left(\frac{k}{k - r(w - C)}\right) & \text{if } w < \frac{k}{r} + C \\ +\infty & \text{otherwise} \end{cases} \tag{6}$$

The date-zero value of the solo entrepreneur's decision problem now is

$$\begin{aligned} & q(\pi, w_0) \\ \equiv\ & \max_T \int_0^{\min(\tau,T)} \left(\pi + w_t - \frac{1-k}{h(t)}\right) e^{-rt} f(t)\, dt + \left(1 - F\left[\min(\tau, T)\right]\right) e^{-r\min(\tau,T)} w_{\min(\tau,T)} \end{aligned} \tag{7}$$

where w_t satisfies (5).

If the entrepreneur drops a project immediately, she ends up with $w - C$. Since q is strictly increasing in π, $\pi^S_{\min}(w)$ uniquely solves

$$q(\pi, w) = w - C. \tag{8}$$

For $w \geq C + k/r$, $\pi^S_{\min}(w)$ is a constant that we shall denote it simply by $\pi^S_{\min}$ (we shall show later that the situation is as depicted in Figure 7). Differentiating (7), the solo entrepreneur's FOC in the region where $T < \tau(w)$ is

$$\begin{aligned}
0 &= \left(\pi + w_T - \frac{1-k}{h(T)}\right) f(T) - f(T)\, w_T - (1 - F[T])\, r w_T + (1 - F[T])(r w_T - k) \\
&= \left(\pi - \frac{1-k}{h(T)}\right) f(T) - (1 - F[T])\, k \\
&= \left(\pi - \frac{1}{h(T)}\right) f(T),
\end{aligned}$$

i.e.,

$$\pi = \frac{1}{h(T^S)}.$$

Therefore, if the value of this problem is positive, then the solution is $\min\left(\tau[w], T^S(\pi)\right)$, otherwise it is zero. Let $\pi_{\min}(w)$ denote the smallest realization of π for which $q(\pi, w) \geq w - C$. That is, $\pi_{\min}(w)$ is the lowest-quality project that an entrepreneur with initial wealth w will be willing to pursue further. Any project quality below $\pi_{\min}(w)$ she would terminate at once. Then the optimal stopping time, T^S, of a solo entrepreneur with initial wealth w is defined as follows

$$T^S(\pi) = \begin{cases} h^{-1}\left(\frac{1}{\pi}\right) & \text{if } \pi > \pi_{\min}(w), \\ 0 & \text{otherwise.} \end{cases} \tag{9}$$

Although w enters its definition, we suppress it in the notation.

We refer now Figure 2 which is the solo entrepreneur's counterpart to the planner's version of the same thing in Figure 1. The decision rule in (9) is similar to the socially optimal rule in (4). One point about (9) should be clarified with the help of the figure: Since $1/\pi_{\min}$ is higher for wealthier entrepreneurs, and so terminations at youngest *strictly positive* ages will be observed among the richest entrepreneurs. But this does not mean that the rich entrepreneurs are less patient. The interval $[\pi_{\min}(w_1), \pi_{\min}(w_2)]$ consists of projects that entrepreneur 2 would terminate right away, but that entrepreneur 1 begins to terminate only at date $T^S(\pi_{\min}[w_1]) > 0$. Conditional on π, however, termination dates are not affected by w, as illustrated by the point $T^S(\pi)$ which does not depend on w.

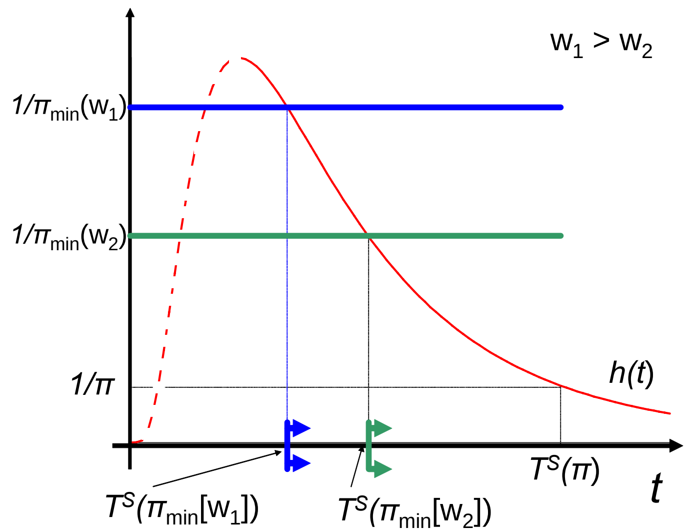


Figure 2: THE DETERMINATION OF $T_w^S(\pi)$ FOR TWO DIFFERENT WEALTH LEVELS

The Socially Optimal Financing Mode

Equations (3) and (9) characterize the optimal decisions on individual projects given the decisions regarding the financing mode. It remained to determine whether an entrepreneur should go solo, seek VC-backing, or invest with a bank. Next, we restrict attention to the question whether an entrepreneur with wealth w should go solo or invest with a bank if VC-backing was not an option. We shall show that there is a cutoff level of wealth, w^*, above which the entrepreneur should go solo and otherwise should invest with a bank. Finally, we characterize those entrepreneurs who should get VC-backing.

The problem of an entrepreneur with limited wealth is that if she goes solo she might run out of money. That is, although it is socially optimal to support a project, a solo entrepreneur is unable to do so because of her wealth constraint. Indeed, the social value of a VC in our model comes from his ability to finance *poor* entrepreneurs. Hence, those entrepreneurs should be matched with VCs who do not have enough liquidity to finance their own projects for long enough time. We shall assume that there are many poor entrepreneurs, with wealth below w^*, who would invest with a bank instead of going solo in the absence of VCs. Then, in the socially optimal outcome VCs are backing only (some of the) entrepreneurs that have wealth less than w^*.

Going Solo vs. Investing with a Bank. —From (6), $\tau'(w) = k - r[w - C]/k$. Differentiating in (7), with respect to w yields

$$\frac{\partial q(\pi,w)}{\partial w}=\begin{cases} 1+\left(\frac{k-r[w-C]}{k}\right)\left(\pi-\frac{1}{h(\tau[w])}\right)e^{-rt}f(\tau[w]) & \text{if } \tau(w)<T^{S}(\pi) \\ 1 \quad \text{otherwise.} & \end{cases} \quad . \tag{10}$$

The expected social value of a solo entrepreneur with wealth w is

$$Q^{S}(w)=\int q(\pi,w)\,dG(\pi).$$

Lemma 1 *For $w<k/r+C$,*

$$\frac{\partial Q^{S}}{\partial w}>1.$$

The intuition behind the statement of this lemma is the following. A budget-constrained entrepreneur can use an additional dollar to prolong the time of supporting her project, instead of using it for consumption. The marginal value of consumption would be exactly one. Since sometimes it is socially efficient to finance the project longer than the budget-constrained entrepreneur can afford, her marginal value for a dollar exceeds one.

Proof. By (10), $dQ^{S}/dw\geq 1$, and it is strictly greater than unity whenever there are at least some realizations of π such that w is not enough to support the project up to the socially optimal time. But $T^{S}(\pi)$ is unbounded if π is. Therefore, $dQ^{S}/dw>1$ whenever $w<k/r+C$. ■

On the other hand, if $w\geq k/r+C$, the entrepreneur can finance her project indefinitely if she wants. Since by assumption the expected social value of a project is positive, $Q^{S}(w)>w$ whenever $w>k/r+C$. Indeed, we have

Lemma 2 *For $w\geq\frac{k}{r}+C$,*

$$Q^{S}(w)=w+\sigma \qquad \textit{where} \quad \sigma\geq W\left(1-E_{\pi,t}e^{-r\min(t,T^{*}(\pi))}\right)\geq 0. \tag{11}$$

An entrepreneur with $w>k/r+C$ can already support her project as long as it is socially optimal. She would use an additional dollar for consumption. Hence, her value for an additional dollar is exactly one, explaining why $Q^{S}(w)=w+\sigma$.

Proof. A rich-enough entrepreneur generates surplus

$$\begin{aligned} \sigma &= -C+\int_{0}^{T^{S}(\pi)}\left(\pi-\frac{1}{h(t)}\right)e^{-rt}f(t)\,dt \\ &= -C+E_{\pi}\left\{\int_{0}^{T^{S}(\pi)}e^{-rt}\left(\pi+W-\frac{1}{h(t)}\right)f(t)\,dt-W\int_{0}^{T^{S}(\pi)}e^{-rt}f(t)\,dt\right\} \\ &\geq -C+E_{\pi}\left\{\int_{0}^{T^{*}(\pi)}e^{-rt}\left(\pi+W-\frac{1}{h(t)}\right)f(t)\,dt-W\int_{0}^{T^{*}(\pi)}e^{-rt}f(t)\,dt\right\}. \end{aligned}$$

The second equality holds because we just added and subtracted $W\int_0^{T^S(\pi)} e^{-rt} f(t)\,dt$. The inequality holds because although $T^*(\pi)$ is a feasible policy for the entrepreneur, $T^S(\pi)$ is the optimal one. But, as we show later in (13), W is defined by the following equation

$$W = -C + E_\pi \left\{ \int_0^{T^*(\pi)} \left(\pi + W - \frac{1}{h(t)} \right) e^{-rt} f(t)\,dt + W e^{-rT^*(\pi)} \left(1 - F\left[T^*(\pi)\right]\right) \right\}.$$

Therefore

$$\begin{aligned}
\sigma &\geq W - W E_\pi \left\{ e^{-rT^*(\pi)} \left(1 - F\left[T^*(\pi)\right]\right) + \int_0^{T^*(\pi)} e^{-rt} f(t)\,dt \right\} \\
& \quad W - W E_\pi \left\{ \int_{T^*(\pi)}^{\infty} e^{-rT^*(\pi)} f(t)\,dt + \int_0^{T^*(\pi)} e^{-rt} f(t)\,dt \right\} \\
&= W - W E_\pi \int_0^{\infty} e^{-r\min(t,T^*(\pi))} f(t)\,dt = W\left(1 - E_{\pi,t} e^{-r\min(t,T^*(\pi))}\right).
\end{aligned}$$

■

The marginal solo entrepreneur.—Thus we have shown that $Q^S(w)$ must look as drawn in Figure 3. It starts from zero when $w = C$ because at $w = C$, right after paying C, the entrepreneur would have no money left to continue supporting the project; thus $Q^S(C) = 0$. As w reaches $k/r + C$, $Q^S(w)$ reaches $w + \sigma$ which, in the case where $T^*(\pi) > 0$ for some π is strictly above the 45^0 line. We shall argue that from the Intermediate Value Theorem it follows that there exists a unique value of wealth, denoted by w^*, that solves the equation

$$Q^S(w) = w. \tag{12}$$

Thus w^* is the wealth of the poorest solo entrepreneur. Figure 3 depicts the choice between going solo and investing with a bank and the determination of w^*. The payoff, $Q^S(w)$ is continuous in w and is not defined if $w < C$ because the entrepreneur cannot pay the cost C. That w^* is unique follows because by Lemma 1 $\partial Q^S/\partial w > 1$ for $w < k/r + C$, and because $Q^S\left(\frac{k}{r} + C\right) > k/r + C$. This latter inequality holds, because the social value of a project is strictly positive. Therefore, at the point where the Q^S curve intersect with the 45-degree line, the slope of Q^S strictly exceeds unity. (Recall from Lemma 1 and Lemma 2 that the slope of Q^S turns into one only at $k/r + C$.)

Who Should get VC-backing and the Value of a free VC—We turn to the determination of W. We maintain the assumption that VCs finance those entrepreneurs who would otherwise not go solo but with invest with a bank. (Later, we provide a condition on the wealth distribution of the entrepreneurs which guarantees that this

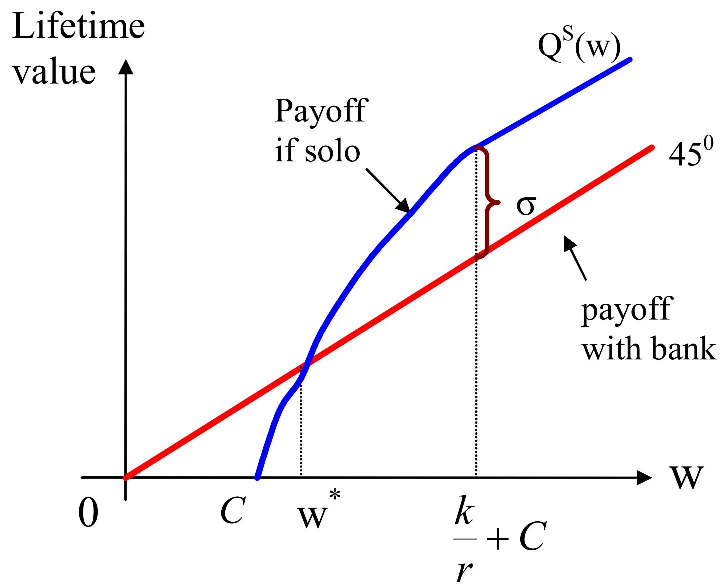


Figure 3: The determination of w^*

is indeed socially optimal.) Hence, the social value of a free VC is determined by the following equation

$$\begin{aligned} W &= -C + \int_0^{\infty} \int_0^{T^*(\pi)} \left(\pi + W - \frac{1}{h(t)} \right) e^{-rt} f(t)\, dt dG(\pi) \qquad (13) \\ &\quad + W \int e^{-rT^*(\pi)} \left(1 - F\left[T^*(\pi)\right]\right) dG(\pi) . \end{aligned}$$

From this

$$W = \frac{-C + \int_0^{\infty} \int_0^{T^*(\pi)} \left(\pi - \frac{1}{h(t)} \right) e^{-rt} f(t)\, dt dG(\pi)}{1 - \int_0^{\infty} \int_0^{\infty} \max\left(e^{-rt}, e^{-rT^*(\pi)}\right) f(t)\, dt dG(\pi)}. \qquad (14)$$

To see that (14) indeed uniquely defines W, note that when W is zero, the right-hand side is positive. This is because if $W = 0$, $T^* = T^S$ and the social value of a project is positive. If W goes to infinity the right-hand side becomes negative because T^* converges to zero, and hence the $-C$ part will dominate. Finally, since the right-hand side is decreasing and continuous in W, the existence of unique solution is guaranteed by the Intermediate Value Theorem.

The average duration of the average venture-backed project, $\bar{t}$, can be computed as follows

$$\bar{t} = \int \int_0^{\infty} \min\left(t, T^*[\pi]\right) f(t)\, dt dG(\pi) .$$

At any point in time, there is a measure $x/\bar{t}$ of free VCs. Recall that at any instance of time there is an inflow of λ of new entrepreneurs. Among them there is a measure of $\lambda\Psi(w^*)$ who has so little wealth that, in the absence of VCs, would choose to invest with a bank. If $x/\bar{t} \leq \lambda\Psi(w^*)$ then it is indeed socially optimal to match the VCs with these entrepreneurs for whom it would otherwise not be socially optimal to go solo.

The following proposition summarizes our findings:

Proposition 1 *If $\lambda\Psi(w^*) > x/\bar{t}$, the socially optimal outcome is defined as follows:*

(i) An entrepreneur with initial wealth $w > w^*$ goes solo. A measure of $x/\bar{t}$ entrepreneur gets VC-backing at every instance of time, each of them with wealth less than w^*. The rest of them invest with a bank.

(ii) The termination decision of a venture-backed project is determined by (3), and that of the solo project is by (9).

3.2 The Competitive Outcome

In what follows we show that under some conditions on the distribution, Ψ, of entrepreneurs' wealth, the socially optimal outcome is implemented as a competitive equilibrium. Recall, that the social value of a free VC, W, plays an important role in determining the socially optimal decisions. In order that the VC makes optimal decisions, it is essential that his market value (when he is free) should be exactly W. But this implies that the VC must have enough market power to be able to extract all the surplus from individual projects. We guarantee this market power to the VCs by assuming that there are more entrepreneurs who is willing to seek VC-backing than VCs. In other words, we assume that $\lambda\Psi(w^*)$ is large enough compared to the available free VCs, $x/\bar{t}$.

But this is not the whole story. Recall, that in our model there is a double-sided moral hazard problem at work. Neither the effort of an entrepreneur, nor the investment of a VC is contractible. Hence, the VC must be able to provide a contract to the entrepreneur which induces the socially efficient termination decisions by both parties and, in addition, enables the VC to extract the whole surplus.

Recall, a contract consists of two numbers (p, s), where p is paid by the VC to the entrepreneur before π is realized, and s is the sharing rule upon the realization of π. We shall show that if the sharing rule is

$$s^* = \frac{a}{1 + rW},$$

the termination rules of both parties are indeed the socially optimal ones. But how can the VC extract the whole surplus from the entrepreneur?

Let $Q^{VC}(s)$ denote the continuation value *to the entrepreneur* from a contract specifying sharing rule s, conditional on both parties supporting the project up to

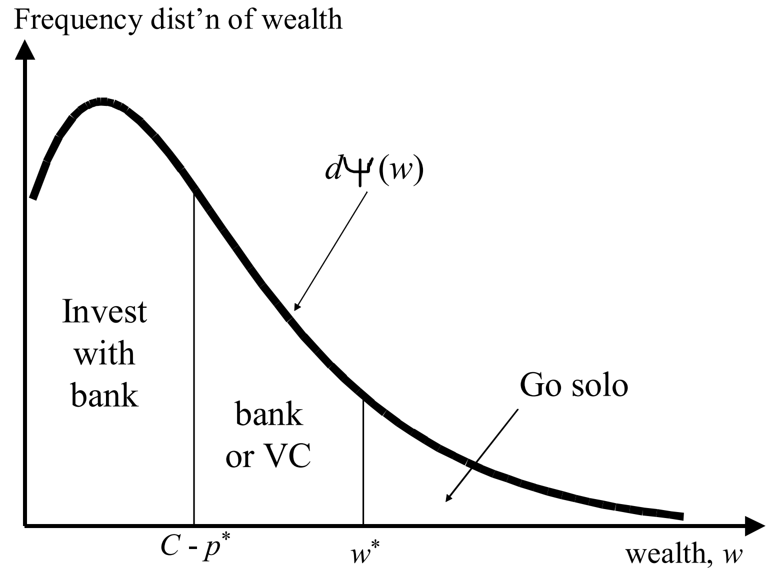


Figure 4: The equilibrium allocation of entrepreneurs to activities

$T^*(\pi)$. Then

$$Q^{VC}(s) = \int \left\{ \int_0^{T^*(\pi)} \left(s\pi - \frac{a}{h(t)} \right) e^{-rt} f(t)\, dt \right\} dG(\pi).$$

The determination of p conditional on s.—The portion of C that the entrepreneur must finance is $C - p$. The VC extracts all the rents if the entrepreneur extracts none. Then

$$p = C - Q^{VC}(s). \tag{15}$$

As long as $Q^{VC}(s) > 0$, the VC does not pay the entire fixed cost. The entrepreneur must pay $C - p$ up front, and some potential entrepreneurs will have wealth insufficient to cover this amount. These are people with wealth below $C - p$ in Figure 4.

The selection of entrepreneurs into activities.—Entrepreneurs' choices of the mode of investment are described in Figure 4. The fraction of entrepreneurs that wishes to get VC backing is $\Psi(w^*)$. But of these, the fraction that can also afford to pay an up-front product-development cost of $C - p$ is just $\Psi(w^*) - \Psi(C - p)$. This is the area "bank or VC" in Figure 4. Hence the distributional assumption we need is

$$\Psi(w^*) - \Psi(C - p^*) > \frac{x}{\lambda \bar{t}}. \tag{16}$$

Theorem 1 *If (16) holds, the socially optimal outcome is also a competitive equilibrium outcome supported by the following strategies:*

(i) A VC always offers the contract (p^*, s^*). *If the contract is accepted, he follows the socially optimal decisions, defined by (3).*

(ii) An entrepreneur with wealth $w \geq w^*$ *goes solo, and follows the socially optimal termination rule defined by (9).*

(iii) An entrepreneur with wealth $w \in (p^*, w^*)$ *seeks VC-backing with probability* $x/(\bar{t}\lambda(\Psi(w^*) - \Psi(C - p^*)))$ *and invests with a bank otherwise. Entrepreneurs seeking VC backing accept the contract offered by the VC, and follow the socially optimal decisions defined by (3).*

(iv) An entrepreneur with wealth $w \leq C - p^*$ *invests her money with a bank.*

Notice that (16) requires that w^* be larger than $C - p^*$. This turns out to be true because we have the following two Lemmas:

Lemma 3

$$p = \frac{k + rW}{1 + rW} C \tag{17}$$

Proof. The proof is contained in the seven lines preceding eq. (50) of the Appendix. ■

Lemma 4

$$w^* > C - p^*. \tag{18}$$

Proof. .We have $Q^S(C) = 0$ because paying C, the entrepreneur would have no money left to continue supporting the project. Using (12), we then have $w^* > C$. But since $k \leq 1$, $p \leq C$, from which (18) follows. ■

Notice that in the equilibrium described above, the VCs extracts all the surplus from individual projects.[8] If (16) is true, then (i) there are enough *poor* entrepreneurs who prefer to go with a VC and, (ii) among these there are enough who have enough wealth to finance their share of C.. Since the VCs extracts all the social surplus from the projects, their market value will be exactly the social value of a VC, W.

The equilibrium is further described in Figure 5. The Figure takes the equilibrium features of the contract as given, except for p. That is, as p varies, s is held fixed at s^*. The vertical axis of Figure 5 measures the up-front cost to the entrepreneur. The Figure may be explained as follows:

1. If there were no VCs, a total of $\Psi(w^*)$ entrepreneurs would simply abandon their projects and invest their wealth with banks, and the remaining $1 - \Psi(w^*)$ would go solo as shown in Figure 3.

[8] Moreover, once terminated by a VC, the entrepreneur would not wish to continue the project alone (either through self finance or bank finance) because the VC retains his equity in the project even after ceasing to invest in it. Thus the entrepreneur's reward would not rise, but her costs would, and so she would strictly prefer to stop right away.

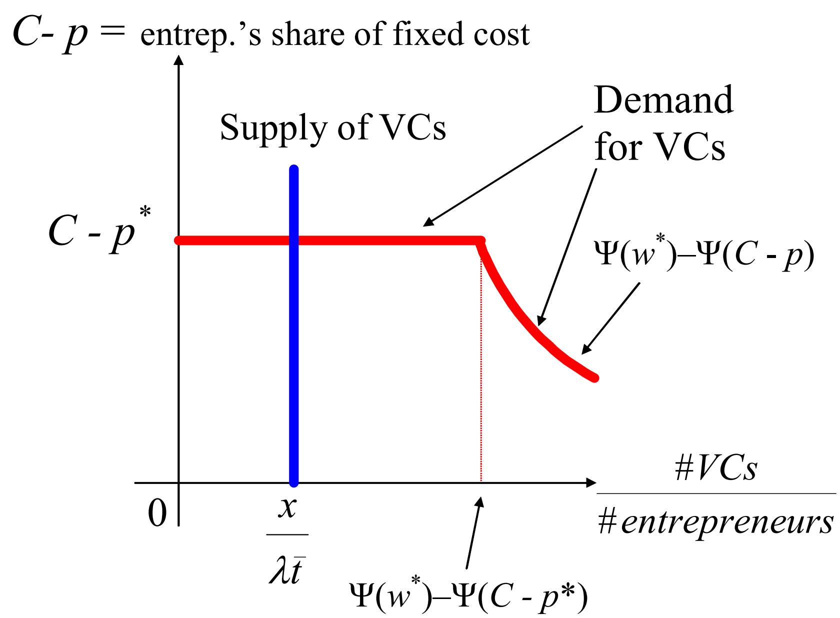


Figure 5: THE DETERMINATION OF p WHEN $s = s^*$.

2. Since investing with a bank offers the entrepreneur zero rents, the entrepreneurs' demand for VCs is infinitely elastic at $C - p^*$ up to the point $\Psi(w^*) - \Psi(C - p^*)$. The poorest $\Psi(C - p^*)$ entrepreneurs could not afford the up-front cost.

3. At an up-font cost any higher than $C - p^*$, no one would demand VC services. At any cost below this value, the payoff to going with a VC would strictly dominate that of going to a bank. But not all $\Psi(w^*)$ of the entrepreneurs could afford to sign with a VC; at p^* the demand curve has a kink; as p rises above p^* there is a gradual rise in the number of entrepreneurs that can afford to develop their project *and* that are willing to sign with a VC.

3.3 Proof of Theorem 1

First, we prove that given the decision about the financing mode, the entrepreneurs' as well as the VCs' decisions regarding the termination time of a project are indeed socially optimal. That is, we prove the second part of claims (i), (ii), and (iii) of Theorem 1. If an entrepreneur decides to go solo, then she is the one who incurs all the costs related to the project, but she also enjoys all the potential benefits. In other words her costs and benefits are identical to the social costs and benefits, hence she obviously follows the socially optimal decision rules describe by (9). Therefore, we only have to show that if a project is venture backed, the entrepreneur and the VC both follow the socially optimal decision rule defined by (3).

Second, we show that given the decisions regarding the individual projects, the decisions regarding the financing mode are as described in the first parts of claims (ii), (iii), and (iv) of the theorem.

Since the VC extracts all the surplus, he obviously has no incentive to offer a different contract.

Incentive Compatibility of the Contract (p^*, s^*)

We analyze now the incentives of the agents to support the project after a contract (p, s) is signed and both parties learn the value of π.

Entrepreneur.—Suppose first, that the entrepreneur trusts that the project is always financed by the VC, and that she will get $s\pi$ if the project is successful. Since the project has no salvage value, if it is terminated the entrepreneur gets zero as her terminal payoff. Recall, new ideas occur only to new entrepreneurs. Therefore she solves

$$V^E(\pi) \equiv \max_T \int_0^T \left(s\pi - \frac{a}{h(t)}\right) e^{-rt} f(t)\, dt. \tag{19}$$

If the solution, $T^e(\pi)$, is interior it is defined by the corresponding first-order condition:

$$h\left(T^E(\pi)\right) = \frac{a}{s\pi}. \tag{20}$$

The local second-order condition, which is also is also the sufficient condition, is again $h'(T^e(\pi)) < 0$. Finally, if the value of the maximization problem is negative, she does not start to exert effort.

VC.— Recall, the market value of a free VC, that is the expected payoff of a VC who is not yet in a contractual relationship with an entrepreneur is just W. Suppose now, that the VC trusts that the project is always supported by the entrepreneur, and he gets $(1-s)\pi$, if the project succeeds. The VC's maximization problem after signing the contract is

$$V^{VC}(\pi) = \max_T \int_0^T \left((1-s)\pi + W - \frac{k}{h(t)}\right) e^{-rt} f(t)\, dt + e^{-rT}(1 - F[T])\, W. \tag{21}$$

In other words, the VC can find a new project immediately after one is over (whether it was terminated or whether it succeeded). If the solution, $T^{VC}(\pi)$, is interior, it must solves the first-order condition

$$h\left(T^{VC}(\pi)\right) = \frac{k + rW}{(1-s)\pi}. \tag{22}$$

The sufficient condition is again $h'\left(T^{VC}(\pi)\right) < 0$. If the value of the maximization problem in (21) is less than W, the VC does not invest in the project.

Incentive compatibility.— The agents stop supporting the project at the same moment if and only if $h\left(T^{VC}(\pi)\right) = h\left(T^{E}(\pi)\right)$. From (20) and (22) it follows that this equality holds if and only if

$$\frac{a}{s\pi} = \frac{k + rW}{(1-s)\,\pi}.$$

But this requires that

$$s = \frac{a}{(a + k + rW)}. \tag{23}$$

Optimality.— Recall from (3) that the socially optimal termination decision, $T^*(\pi)$, satisfies $h(T^*(\pi)) = (1 + rW)/\pi$. Hence, in order to achieve the socially optimal rule, we need that

$$\frac{a}{s\pi} = \frac{1 + rW}{\pi}.$$

Given the incentive-compatible s in (23), we need that

$$a + k + rW = 1 + rW.$$

But this is true since $a + k = 1$.

It remained to show that, if s is defined by (23), the minimum value of π which makes the RHS of (1) at least W, i.e., $\pi_{\min}$, is the same as the value of π which makes the RHS (21) at least W, and that the RHS of (19) is nonnegative. That is

Lemma 5 *Let* $\pi_{\min}$ *solve (2). If* s *is defined by (23), then* (*i*) $V^{VC}(\pi_{\min}) = 0$ *and* (*ii*) $V^{E}(\pi) = 0$

(proved in the Appendix).

Since V^{VC} and V^{E} are increasing in π, the Lemma implies that both are nonnegative for all $\pi \geq \pi_{\min}$. Thus we have shown that if $s = a/(a + k + rW)$, then both (20) and (22) become just (3). That is, for all π, $T^{E}(\pi) = T^{VC}(\pi) = T^*(\pi)$. This implies that the VC as well as the entrepreneur support the project up until it is socially optimal to support it. This shows the second parts of claims (*i*) and (*iii*) of the Theorem, to the effect that both parties follow the socially-optimal termination decisions defined in (3).

The Choice of Financing Mode of an Entrepreneur

We now show the first parts of claims (*ii*) and (*iii*) and claim (*iv*) of Theorem 1.

Suppose first, that an entrepreneur has initial wealth $w \leq C - p^*$. Since $w < w^*$, she is better off putting her money into the bank instead of going solo. Furthermore, she cannot contract with a VC, because she does not have enough liquidity to finance $C - p^*$ up-front. Hence, she invests with the bank.

Suppose that $w \in (C - p^*, w^*)$. Since $w < w^*$, entrepreneur is still better off by investing with a bank instead of going solo. However, she has enough wealth to finance $C - p^*$ of the fixed cost. Since the VCs extract all the surplus from the projects, these entrepreneurs are indifferent between seeking VC-backing or investing with a bank. So they can randomize according the claim (iii) of Theorem 1.

If $w > w^*$, the entrepreneur will go solo, since $Q^S(w) > w$, and her other options all provide her with a payoff of w.

Discussion of the efficiency result

The fact that the equilibrium contract induces the socially optimal stopping time may at first be surprising, because there is a two-sided moral hazard problem in our model. Notice however, that if $s = a/(a + k + rW)$, then the objective function of the entrepreneur is simply a times the objective function of the social planner. True enough, when the entrepreneur solves her maximization problem, she only cares about her own cost, a, instead of the social cost $a + k$. However, if $s = a/(a + k + rW)$, then the entrepreneur cares only about her own benefit, $[a/(a + k + rW)]\pi$, instead of the social benefit π. Both the cost and the benefit in the maximization problem of the entrepreneur are down-scaled by a compared to the social surplus function. Therefore they are maximized at the same value of T.

Two assumptions guarantee that the socially optimal outcome is a competitive equilibrium. First, there must be few VCs relative to the number of entrepreneurs that seek VC-backing. And, second, among these entrepreneurs there must be sufficiently many that have enough liquidity to pay $Q^{VC}(s^*)$ up-front.

The first assumption is crucial to our result. This assumption provides the VCs with market power. They are able to offer contracts that enable them to extract the full social surplus. That is why the market value and the social value of a free VC are the same.

The assumption regarding the number of rich entrepreneurs among those who would not go solo is far less important. More complicated contracts would make it possible to extract surplus from entrepreneurs who do not have enough cash in hand to start with. Recall, that we have restricted attention to contracts which specify a time-independent sharing rule, that is s cannot depend on the time when the project succeed.[9] The VCs could extract surplus from a more liquidity constrained entrepreneur by offering contracts when this sharing rule is increasing. Recall that with the fixed s the entrepreneur was only indifferent between exerting effort and shirking at the time of termination, but strictly preferred to exert effort anytime before. If s was allowed to change over time, the entrepreneur could have been made indifferent between working and shirking at any time before the termination of the project, and by such contracts surplus could have been extracted from *poor* entrepreneurs too.

[9] In reality the share of the entrepreneur falls as the project ages. Our model would generate a declining s if the ratio $\frac{a}{a+k}$ were to decline with t.

Evidence on the efficiency of equilibrium

Efficiency holds only if (16) holds, i.e., only if in Figure 5 the supply is to the left of the kink in the demand curve. When this condition holds, small shifts in the supply curve leave the equilibrium return to venture capital unchanged. Kaplan and Schoar (2005, Table 13) estimate the effect that entry of new funds has on venture fund returns as a whole. Such entry would represent a shift to the right of the supply curve in Figure 5. They also control for the returns on the Nasdaq Composite Index in the year a fund was started because, as they also find in Table 12, that funds are more likely to get started in boom years. They find that the correlation between fund returns and the logarithm of the number of new entrants is negative but statistically insignificant. In some variations, the returns on the new funds themselves decline, but the returns on the established funds do not. Overall, their results support the view that the intersection in Figure 5 occurs at the point where the demand curve is flat.

On the other hand, Gompers and Lerner (2000) find that outward shifts in the supply of venture capital act to raise the total amount that VCs pay into the companies they oversee. In our model this can happen only if such supply shifts also lower the rate of discount, r. If r fell, $T = \frac{\rho}{1+rW}$ would rise, and the total amount of investment per portfolio company would also rise. Similarly, Hochberg, Ljungqvist, and Lu (forthcoming) find that the likelihood of getting to IPO rises following a positive shock to the supply of funds. These results suggest that the intersection in Figure 5 occurs at the point where the demand curve slopes down.

Discussion of work on efficiency in the market for venture capital

Bergemann and Hege (1998, 2005) argue that dynamic contracts between entrepreneurs and VCs are inefficient relative to first best. In their model the project succeeds with some probability in each period, and the payoff is proportional to the invested funds. As in our setup, as time passes without success, agents down-date their prior, that is, they become more and more pessimistic. The main difference is that Bergemann and Hege (1998) focus on the following moral hazard problem: The entrepreneur can divert the invested funds to private consumption (with or) without the VC observing it. The trade-off the entrepreneur then faces is the following: On the downside, if she diverts the funds, she reduces the probability that the project succeeds. On the upside, if she diverts the funds: (1) she benefits directly by consuming them, (2) she potentially prolongs the time that she gets the stream of funds. The main result of Bergemann and Hege (1998) is that the optimal contract specifies a decreasing stream of funds. The project is supported for a time that is shorter than would be socially efficient, and the project gets less funds. The reason stems from the trade-off described above. The investment stream should be specified such that the entrepreneur has no incentive to divert it. If it is decreasing, the entrepreneur understands that if she diverts it, then: (1) the payoff upon success decreases (recall it is proportional to the size of the funds), and (2) the future stream of funds is less attractive, because it is decreasing.[10]

Inderst and Muller (2004) and Michelacci and Suarez (2004) build search models and assume that Nash bargaining divides the rents between the VC and the entrepreneur. As is the case in search models with a matching function, the "Hosios condition" (which states factor shares in the constant-returns-to-scale matching function should equal the factors' relative bargaining strength) must hold in order that the equilibrium be efficient. It is pure coincidence if that equality should obtain, and so generically these models imply inefficiency of equilibrium – policies that change incentives for entry by one side or the other can generally improve the sum of the payoffs. In our model, by contrast, efficiency holds on an open set of all parameter values; although (16) is not in terms of primitives, it is seen that w^*, p^0, and $\bar{t}$ are continuous in the parameters of the model, and, hence, that there is a range of all parameters for which the condition holds.

In a sense the Hosios condition does hold in our model: We have a constant return to scale matching function that is Leontieff, with VCs always on the short side. This implies that the elasticity of the matching function with respect to the stock of venture capital is one. Moreover, VCs get all the bargaining power, and so in this sense the Hosios condition holds.

Our model takes the relative numbers of entrepreneurs and VCs as exogenous. Since VCs get the full social value of their capital, if we endogenized venture capital we would expect that an optimal amount of it would be created. Entrepreneurs receive a zero return on their ideas. Indeed, the marginal social value of another entrepreneur with an idea is zero – because society does not have a free VC to finance her idea, the entrepreneur would generate as much social benefit if she were invest her wealth with a bank. Figure 5 shows, however, that the poorest $\Psi(w^*)$ entrepreneurs receive no value from their ideas. If getting an idea was costly, they would have no incentive to pay such a cost. In that case, optimality would survive only if entrepreneurs received ideas about projects incidentally, say through learning by doing, or if they were born with ideas.

[10] Lerner (1998) has argued that Bergemann and Hege's assumption that the manager can divert funds is unrealistic because the VC usually monitors activities in his firms on a weekly and sometimes on a daily basis, attends monthly board meetings and so forth.

4 Empirical implications

This section lists some qualitative implications of the model and compares them with evidence. The next section presents estimates of the model.

4.0.1 Good projects receive more investment rounds

Gompers (1995) finds that bad projects tend to be identified early and get dropped, and that it is the good projects that receive more investment. This happens in our model: The amount that the VC expects to invest is increasing in π. First of all, projects with $\pi \leq \pi_{\min}$ receive no investment beyond the initial outlay C. For a project with $\pi > \pi_{\min}$, investment proceeds for $T^{VC}(\pi)$ rounds, a number that solves the equation

$$h(T) = \frac{1 + rW}{\pi}. \tag{24}$$

At the point of intersection $h' < 0$, as shown by the solid line in Figure 1. The dashed portion is not admissible because the second-order conditions fail. When the solution exists,

$$\frac{\partial T}{\partial \pi} = -\frac{h(T)}{h'(T)} > 0,$$

so that the maximum number of investment stages rises with the project's quality. We now illustrate this in Figure 6. Until date $t = h^{-1}([1 + rW]/\pi_{\min})$, no projects are being terminated, and successes are drawn from the distribution $G(\pi \mid \pi \geq \pi_{\min})$. Projects to the left of $\pi_{\min}$ are terminated right away. At $t = h^{-1}([1 + rW]/\pi_{\min})$, the truncation point, $(1 + rW)/h(t)$ starts to move to the right. Thus the conditional mean of the projects that are funded rises. Let $\Gamma_t(\pi)$ be the distribution of π among projects that bear fruit at date t. Then for t larger than the value at which the mode of h occurs (so that the condition $h(t) = \frac{1+rW}{\pi}$ represents a maximum),

$$\Gamma_t(\pi) = \frac{G(\pi) - G\left(\max\left\{\pi_{\min}, \frac{1+rW}{h(t)}\right\}\right)}{1 - G\left(\max\left\{\pi_{\min}, \frac{1+rW}{h(t)}\right\}\right)} \tag{25}$$

for $\pi \geq \max\left\{\pi_{\min}, \frac{1+rW}{h(t)}\right\}$. For t below the mode of h,

$$\Gamma_t(\pi) = \frac{G(\pi) - G(\pi_{\min})}{1 - G(\pi_{\min})} \tag{26}$$

for $\pi \geq \pi_{\min}$.

This is where the assumption that π and τ are independent has bite. A sufficient negative correlation between π and τ would overturn the result. If all high-π projects had low τ's, and if the low-π projects had high τ's but were still worth supporting, the *bad* projects would receive more investment rounds.

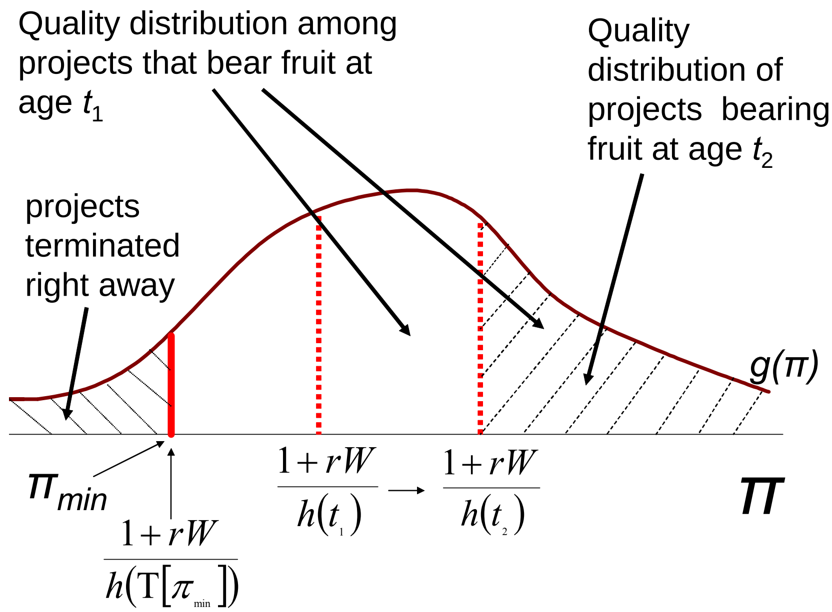


Figure 6: Good projects receive more investment rounds

4.0.2 The excess rate of return to venture capital

It is probable that one can obtain a more reliable estimate of returns to VC from venture-equity fund returns than one can from the returns on individual deals. The data that we have on the latter are known to miss about 15 percent of the investment rounds, and to have selection biases (Kaplan, Sensoy and Strömberg 2002).

The lifetime value of venture capital is W, and the lifetime investment of a VC is C_{PV}. Therefore the lifetime value per lifetime dollar invested is W/C_{PV}. This is also the flow value per period. Thus the excess rate of return to venture capital is

$$\frac{W}{C_{PV}} \equiv \alpha.. \tag{27}$$

To arrive at W and C_{PV} we discount by r, the rate of return required given the risk characteristics of the income stream that the VC faces. That rate depends partly on the covariance of the VC's income stream with the market index, i.e., β. Let r_f denote the risk-free interest rate, and let $r_{S\&P}$ denote the expected return on the market index for which we shall use the S&P 500 as a proxy. Then the CAPM prediction for the expected return on venture capital is

$$r = r_f + \beta\left(r_{S\&P} - r_f\right).$$

To illustrate how (27) works, let time be discrete and assume that each company matures and yields π for sure at the end of one period. Let $p = 1$ and $k = 0$. Then one dollar today yields π dollars a period from now, and so the excess return on the investment would be

$$\alpha = \pi - (1 + r) . \tag{28}$$

Now let us instead calculate the excess return using (27): The VC then invests in a new company every period, and his discount factor is $\frac{1}{1+r}$. Therefore

$$W = \frac{\pi - (1 + r)}{1 - \frac{1}{1+r}} \text{ and } C_{PV} = \frac{1}{1 - \frac{1}{1+r}}.$$

Substituting these values into (27) gives us the same value of α as (28) does.

KS, Cochrane (2004), Ljungqvist and Richardson (2003) and Jones and Rhodes-Kropf (2003) provide estimates of the α's and β's of venture-equity funds. Together with information on r_f and $r_{S\&P}$ this allows us to use the RHS of equation (27) can be calculated and used to constrain the LHS. In reality, VCs are themselves wealth constrained and need outside investors to help them leverage their expertise. The model does not distinguish VCs from venture-fund investors, and so we shall assume that venture-funds returns are the same as the returns that VCs earn on their own investments.

The rate of return on projects that succeed at age t.—Calculating rates if return is easy because the returns all come at the same date t. All that is needed, then, is to bring all costs (which are distributed over $[0, t]$) into date-zero dollars, i.e., to take their present value discounted at the rate r. The VC gets a fraction $(1 - s^*)$ of the payoff, The present value of all costs net of the transfer p^0 would be $\int_0^t ke^{-ru}du + C - p^0$ where p^0 is defined in (15). On a project of quality π, the realized rate of return, $R(t, \pi)$ would solve the equation

$$e^{R(t,\pi)t} = \frac{(1 - s^*)}{\int_0^t ke^{-ru}du + C - p^0}\pi \tag{29}$$

Appendix 4 shows that R is increasing in π and decreasing in t. Now π differs over projects that succeed at t, and their distribution depends on t, being ever more truncated from the left as shown in Figure 6. Since the π's differ, so do the returns. When collapsing a distribution of returns Two concepts are used in the literature. The *geometric* rate of return, call it $R^G(t)$, given by the formula

$$R^G(t) = \int R(t, \pi)\, d\Gamma_t(\pi) , \tag{30}$$

is just the average of the rates of return.

The rate of return on all projects.—When calculating the VC's rate of return, we are concerned with the rate of return on all projects. To do so, first we hold t fixed:

Of all projects that last exactly t periods, a fraction $\frac{h(t)}{h(t)+\psi(t)}$ succeed, and the rest fail. To compute the rate of return on *all* projects that end at date t, we simply would multiply the RHS of (29) and (54) and then proceed as before. This would lower the estimated rate of return on a project that ends at t roughly by a factor of $\frac{h(t)}{h(t)+\psi(t)}$.

4.0.3 The excess rate of return to entrepreneurship

Ideas too are scarce, and entrepreneurs can draw rents on them if they develop them solo. The model implies that wealthier entrepreneurs should receive higher returns because they can develop their projects up to the optimal stopping time. Let C_{PV}^{w} be the PV of costs on an entrepreneurial project:

$$C_{PV}^{w} = C + \frac{1 - \int e^{-r \min(t,\tau[w])} \left| dS\left(t\right) \right|}{r},$$

where $S^{S}\left(t\right) = \left(1 - F\left[t\right]\right)\left(1 - \Phi^{S}\left[t\right]\right)$, where $\tau\left(w\right)$ is defined in (6) and where Φ^{S} is defined in (37). The rate of return of the entrepreneur in excess of r is

$$\varepsilon\left(w\right) = \frac{Q^{S}\left(w\right) - w}{C_{PV}^{w}}. \tag{31}$$

Both direct and indirect costs are included in the numerator, but only direct costs are in the denominator. The denominator is always strictly positive, because $C_{PV}^{w} \geq C$. At the point w^{*}, where the entrepreneur is indifferent between going solo and investing with a bank or VC, the excess return is zero, i.e., $\varepsilon\left(w^{*}\right) = 0$. Since $\frac{\partial Q^{S}(w)}{\partial w} > 1$, the numerator rises with w, but so does the denominator. It rises with the entrepreneur's level of wealth. The excess return becomes flat at the point $C + k/r$, i.e., the point where the solo entrepreneur ceases to be liquidity constrained in any state of the world, i.e., for any realization of π.

The model predicts a higher present value of profits the rich solo entrepreneurs than for the VC. This is because the VC distorts terminations on current projects in order to get to future projects; i.e., he trades off a lower return on each project against a larger number of projects over time. In other words, a part of W comes from the extensive margin, i.e., future projects. The entrepreneur worries only about the intensive margin. On the other hand, for high w, $C_{PV}^{w} > C_{PV}$ because the rich solo entrepreneur terminates later, and therefore we cannot tell which return is higher.

4.0.4 The J curve of cumulative returns

Let $J\left(t\right)$ be the cumulative income. Then we have the differential equation

$$J'\left(t\right) = \left(-k + h\left(t\right)\left(1 - s^{*}\right)\int_{\frac{1+rW}{h(t)}}^{\infty} \pi d\Gamma_{t}\left(\pi\right)\right) S\left(t\right),$$

and the initial condition $J\left(0\right) = -p$, where Γ is defined in (25) and p is given in (15).

4.0.5 The entrepreneur's stake

When fitting the entrepreneur's share $s^* = \frac{1-k}{1+rW}$ in the firm's equity, we shall use Kaplan and Strömberg's (2003, Table 2) numbers for cash flow rights, i.e., the fraction of a portfolio company's equity value that different investors and management have a claim to. Pooling over all rounds, the mean claim of founders 31.1%, that of VCs is 46.7%, and that of other non-VC investors is 22.2%. Since our model does not include non-VC investors, we constrain s^* to the share of founders in claims other than those of the outside investors. That is we should have s^*

$$s^* \approx \frac{31.1}{31.1 + 46.7} = 0.40. \tag{32}$$

A more recent sample that Kaplan, Sensoy and Strömberg (2006) analyze shows that the founders retain a smaller share, 10% to 19%.

In the model, once the entrepreneur signs the contract with the VC, her share of the project drops from unity to s^*, where it remains until the end. Lerner (1994, Table 5) and Kaplan and Strömberg (2003, Table 8) find that the greatest dilution of the entrepreneur's equity stake occurs in the first financing round. Contrary to the model, however, it appears that s continues to fall as the project ages, though at a decelerating rate. The fall is accompanied by a rise in the number of VCs in the syndicate.

4.0.6 Value at IPO

Hochberg (2004) finds that venture-backed firms are worth more at IPO than non-venture-backed firms. We shall show that this is true if the non-venture-backed firm is managed by a wealthy solo entrepreneur. Consistent with us, the minimum value in Table 1 is smaller for the non-venture-backed sample: proceeds were smaller by a factor of two thirds, and size by a factor of almost four.

The average value of a venture-backed company that succeeds at age t is $E\left(\pi|\ T^{VC}(\pi) \geq t\right)$. Similarly, the average value of a firm a wealthy solo entrepreneurs which succeeds at age t is $E\left(\pi|\ T^{S}(\pi) \geq t\right)$. Notice that both T^{VC} and T^S are increasing in π. Hence, in order to conclude $E\left(\pi|\ T^{VC}(\pi) \geq t\right) \geq E\left(\pi|\ T^{S}(\pi) \geq t\right)$ it is enough to show that $T^{VC}(\pi) \leq T^S(\pi)$. We shall prove it in two steps. We first show that the VC is more selective at age zero, and then that he is more selective at all ages. The claim about age-zero terminations is a claim about $\pi_{\min}$ and is summarized in Figure 7. We state it in the form of a Lemma:

Lemma 6 *(Selectivity at date zero): (i) $\pi^S_{\min}(w)$ is increasing in w and reaching $\pi^S_{\min}$ when $w \geq C + k/r$. (ii) Moreover,*

$$\pi^S_{\min} \leq \pi^{VC}_{\min}. \tag{33}$$

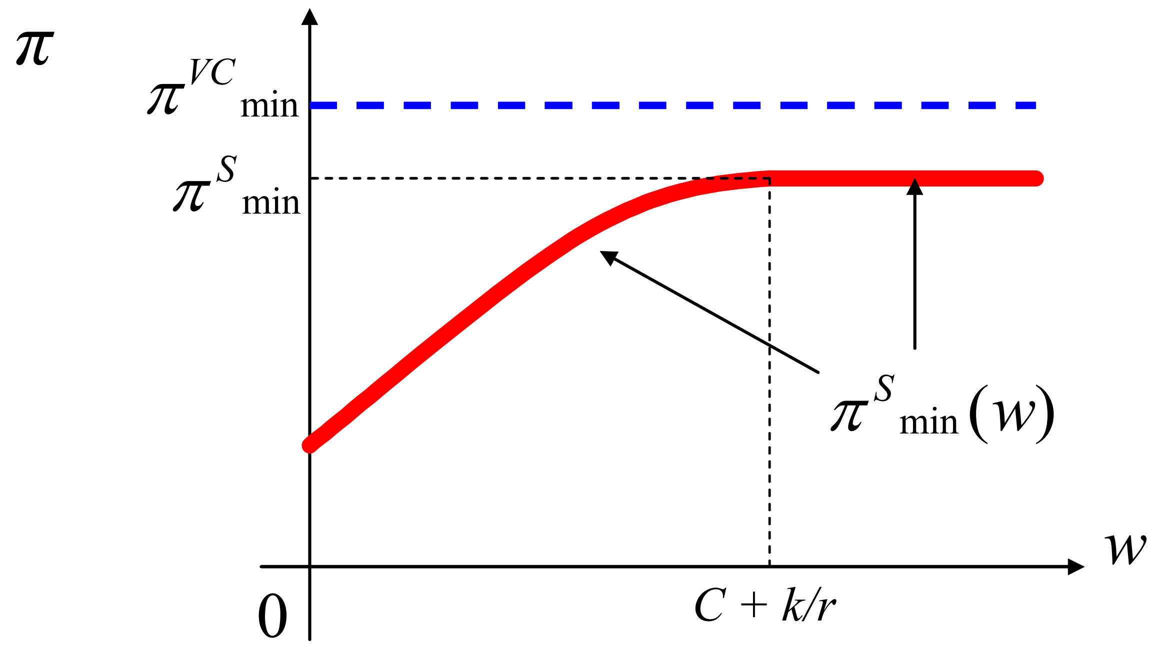


Figure 7: Date-zero terminations by VC and solo firms

Proof. (i) Differentiating in (8), we find that

$$\frac{\partial \pi_{\min}^{S}}{\partial w} = \frac{1 - \frac{\partial q(\pi,w)}{\partial w}}{\frac{\partial q(\pi,w)}{\partial \pi}} \leq 0,$$

because by (10) $\frac{\partial q(\pi,w)}{\partial w} \leq 1$. (ii) We now prove that (33) holds. Recall that for $w \geq C + k/r$, $\pi_{\min}^{S}$ solves

$$\int_{0}^{T^{S}(\pi)} \left(\pi - \frac{1}{h(t)} \right) e^{-rt} f(t)\, dt = 0, \tag{34}$$

and that $\pi_{\min}^{VC}$ solves $V(\pi) = W$. Since V is increasing, it is enough to show that

$V\left(\pi_{\min}^{S}\right)\leq W$. The latter is true because from (1),

$$\begin{aligned} V\left(\pi_{\min}^{S}\right) &= \int_0^{T^{VC}\left(\pi_{\min}^{S}\right)}\left(\pi_{\min}^{S}+W-\frac{1}{h(t)}\right)e^{-rt}f(t)\,dt+e^{-rT^{VC}}\left(1-F\left[T^{VC}\left(\pi_{\min}^{S}\right)\right]\right)W \\ &= \int_0^{T^{VC}\left(\pi_{\min}^{S}\right)}\left(\pi_{\min}^{S}-\frac{1}{h(t)}\right)e^{-rt}f(t)\,dt+\int_0^{\infty}We^{-r\min\left\{t,T^{VC}\left(\pi_{\min}^{S}\right)\right\}}f(t)\,dt \\ &\leq \int_0^{T^{S}\left(\pi_{\min}^{S}\right)}\left(\pi_{\min}^{S}-\frac{1}{h(t)}\right)e^{-rt}f(t)\,dt+\int_0^{\infty}We^{-r\min\left\{t,T^{VC}\left(\pi_{\min}^{S}\right)\right\}}f(t)\,dt \\ &= \int_0^{\infty}We^{-r\min\left\{t,T^{VC}\left(\pi_{\min}^{S}\right)\right\}}f(t)\,dt\leq W. \end{aligned}$$

The first inequality follows from T^S being the solo's optimal termination rule (and not T^{VC}), and the last equality from (34). ■

In Figure 7 the difference $\pi_{\min}^{VC}-\pi_{\min}^{S}$ grows with W. If W were zero, the two lines would coincide.

Proposition 2

$$T^{VC}(\pi)\leq T^{S}(\pi). \tag{35}$$

Proof. Whenever π is such that $T^S(\pi)>0$ (i.e., so that T^S has an interior solution), then h is decreasing at T^S. If $T^{VC}(\pi)>0$, then h is also decreasing at T^{VC} and a comparison of (4) and (3) implies $T^S(\pi)>T^{VC}(\pi)$. Then (35) holds if $T^S(\pi)=0\Longrightarrow T^{VC}(\pi)=0$, i.e., if $\pi_{\min}^{S}\leq\pi_{\min}^{VC}$, which follows from the preceding Lemma. ■

The reason for (35) is the shortage of VCs which gives them market power and a high equilibrium return on investment. Thus the VC's opportunity cost of supporting a project is high, and this makes are impatient with projects that have not yet succeeded. Then (35) also implies that rich solo firms will, on average, reach the stock market later than venture-backed firms, and when they do, their companies will on average be less valuable. We cannot rank terminations of venture-backed firms with those of poor entrepreneurs because they must sometimes terminate early for lack of money.[11]

4.0.7 Terminations

In a sample of 800 venture-backed firms and other startups, Goldfarb, Kirsch, and Miller (2006) and find no significant difference between the failure hazards in the two

[11] The quality of venture capital is homogeneous in our model, hence there is no reason for the VC to try to signal higher ability by taking actions that to outsiders seem successful. "Grandstanding" is said to occur when VC sends companies to an IPO before their time in the hope of establishing a reputation for being able to quickly guide companies to success. A reputable VC can more easily open new funds. Our model does not explain grandstanding, but the finding that $W>0$ is consistent with it.

populations.[12] In our model, Figure 7 implies that initially, at age zero, venture-backed firms must have a higher termination hazard. Thereafter, however, the two cannot generally be ranked. We shall now provide conditions under which the two hazards will be identical at all ages except the initial age, and the example that we shall estimate in the next section will have this property as well.

But we wish to also compare the hazard rates for the two kinds of firms because this is how much of the evidence is presented. We now turn to the definition of these hazards.

The C.D.F.s of terminations.—Let $\Phi(t)$ be the CDF of terminations, and let the density be $\phi(t)$. Now $\pi \sim G(\pi)$ and the low-π projects are terminated first – see Figure 6. By (3), $\pi = (1 + rW)/h\left(T^{VC}\right)$, and so the fraction of projects terminated by date t *conditional on no success* until date t, i.e., $\Pr(T \leq t \mid \tau \geq t)$, is

$$\Phi^{VC}(t, W) = \begin{cases} 0 & \text{for } t = 0 \\ G\left(\pi_{\min}^{VC}(W)\right) & \text{for } t \in \left(0, h^{-1}\left[\frac{1+rW}{\pi_{\min}^{VC}(W)}\right]\right) \\ G\left(\frac{1+rW}{h(t)}\right) & \text{for } t \geq h^{-1}\left(\frac{1+rW}{\pi_{\min}^{VC}(W)}\right). \end{cases} \tag{36}$$

where it is convenient to emphasize the dependence on W of both $\pi_{\min}^{VC}$ and Φ^{VC}. The "rich" entrepreneur with wealth $w \geq C + k/r$ uses the same termination policy as would a VC for whom $W = 0$. Therefore the C.D.F. of terminations for the "rich" solo firm is

$$\Phi^{S}(t) = \Phi^{VC}(t, 0) \tag{37}$$

Since the venture-backed and the solo firm both draw π from the same distribution, (35) implies that for the venture-backed terminations stochastically dominate "rich" solo terminations:

$$\Phi^{VC}(t) \geq \Phi^{S}(t) \tag{38}$$

for all $t \geq 0$.

The *termination hazard* for the venture-backed firms, $\psi^{VC}(t, W) \equiv \frac{\left(\Phi^{VC}\right)'(t)}{1-\Phi^{VC}(t)}$ is

$$\psi^{VC}(t, W) = \begin{cases} \infty & \text{for } t = 0 \\ 0 & \text{for } t \in \left(0, h^{-1}\left[\frac{1+rW}{\pi_{\min}^{VC}(W)}\right]\right) \\ (1 + rW)\,\gamma\left(\frac{1+rW}{h(t)}\right)\left(\frac{-h'(t)}{[h(t)]^2}\right) & \text{for } t \geq h^{-1}\left(\frac{1+rW}{\pi_{\min}^{VC}(W)}\right). \end{cases} \tag{39}$$

where $\gamma(\pi) \equiv \frac{g(\pi)}{1-G(\pi)}$ is the hazard rate of G.

[12] Their "exits" are *firm* exits (and not VC exits which, in the sense often used in the Corporate Finance literature, are successes) se are actually hazards of exiting the primary line of business. Firms may continue in operation doing something that was unrelated to the proposal that the VC looked at. If a company is acquired they treat it as a continuation and not an exit.

The termination hazard of the rich solo firm is

$$\psi^{S}(t) = \psi^{VC}(t, 0). \tag{40}$$

But ψ^{VC} is increasing in W if and only if the function $x\gamma\left(\frac{x}{h(t)}\right)$ is increasing in x, and so this is also when the venture-backed firm will have a higher termination hazard beyond age zero.

Example: Pareto distribution of π.—Let $G(\pi) = 1 - \left(\frac{\pi}{\pi_0}\right)^{-\lambda}$ for $\pi \geq \pi_0 \geq$ so that $\gamma(\pi) = \frac{\lambda}{\pi}$. (This is the functional form we assume for G in the estimated model) Then the expression in the third line of (39) does not depend on W, and for all t for which $\psi^{VC}(t, W)$ and $\psi^{S}(t)$ are positive and finite, we have $\psi^{S}(t) = \psi^{VC}(t, W)$.

4.0.8 Survival to IPO

Ber and Yafeh (2004) find that the probability of survival until the IPO stage is higher for venture-backed companies.[13] (i) Hellmann and Puri (2000) find that venture-backed companies are quicker to market their products. In our model, the outcome is ambiguous. Since $F(t)$ is the same whether a project is venture backed or not, its waiting time to success, $F(t)$, is the same. But because the non-performing venture-backed firms are terminated earlier, there will be fewer venture-backed old firms having an IPO or being acquired. Therefore,

1. Conditional on success, the mean waiting time of venture-backed projects is shorter than that for rich solo projects while
2. Unconditionally, in light of (35), a venture-backed company is more likely to be terminated, and less likely to succeed than a rich solo firm. But the venture-backed firm may be more likely to succeed than a poor solo firm which survives until $\min\left(\tau, \tau[w], T^{S}[\pi]\right)$. On the one hand, $T^{VC}(\pi) < T^{S}(\pi)$, but on the other, $T^{VC}(\pi)$ may exceed $\tau(w)$, i.e., the poor firm runs out of money before its desired stopping time.

Thus, whether venture-backed projects are terminated earlier or later depends on the distribution of wealth. Given that in the population of entrepreneurs w is distributed according to the CDF $\Psi(w)$, we gather these conclusions in a statement about the median firm:

[13] Their control sample, however, their control sample may not correspond to what we have in mind when modeling "solo" entrepreneurs. Ber and Yafeh write: "The control group is a sample of "high-tech companies that were not supported by an Israeli VC fund and raised money from non-VC sources, primarily investment firms that focused on financing high technology but were not organized as VC funds.'

Proposition 3 *If*

$$1 - \Psi\left(\frac{k}{r} + C\right) > \Psi\left(\frac{k}{r} + C\right) - \Psi\left(w^*\right), \tag{41}$$

the median venture-backed firm is likely to (i) *be terminated more quickly,* (ii) *be worth more at IPO than the median solo firm and* (iii) *succeed more quickly conditional on not being terminated.*

Proof. (i) and (ii) follow because when (41) holds, the median w satisfies the conditions of Proposition 3. (iii): For the same reason, for the median solo firm, by (35), $T^{VC}(\pi) \leq T^S(\pi)$ for each π. If it is going to succeed, then, a venture-backed project π must do so earlier than a solo project. And since the ex-ante distribution of π, $G(\pi)$ is the same for venture-backed and solo firms, venture-backed successes are quicker than solo successes. ■

5 Estimating the model

The estimation uses only a venture-backed sample and therefore henceforth we drop the superscript VC when possible. Before getting to the example, we derive the distribution of termination times and the distribution of contract durations or the "Survivor Function."

Contract duration.—The survival of contracts requires that neither a success nor a termination has taken place. Let t denote the date of the "event" that the firm experiences. The event is either a success or a failure, but not both. Only one event per firm can occur. For some firms no event occurs and these are called the "survivors." That is, if τ is date of success and $T(\pi)$ is the date of termination, then $t = \min(\tau, T[\pi])$. Since τ and π are independent random variables, the CDF of t is $1 - S(t)$, where

$$S(t) = (1 - F[t])(1 - \Phi[t]) \tag{42}$$

is the *Survivor function* – the fraction of firms surviving past age t.

The data.—Our data (described in the Appendix) include a distribution of T (terminations) and the distribution of τ (successes) for about 1400 firms, and data on internal rates of return for VC's by age of project completion. We also use information on the VCs' rate of return by age of completed project.

5.0.9 Example: Pareto F and G

We now estimate a five-parameter example which leads to simple formulas, some of which are derived in the Appendix. One object is to estimate W and based on that estimate we shall derive the excess rate of return on venture capital. We note that we shall *not* truncate the waiting time distribution at 10 or 12 years when the venture fund closes. We shall assume that the VC maintains his interest in a company beyond the fund-closing date, and that he can continue to fund it and collect on any return that it generates. This is a important part of the portfolio, containing more than a quarter ($\frac{365}{1355} = 0.27$) of all the firms.

Fitting investment flows.—We shall fit C and k to the data on the investment profile, i.e., the sequences of investment rounds, but converted to flows of investment as a function of time. The sizes of the rounds are reported in Appendix Table A2. The conversion procedure is also described there.

Fitting the success hazard, h.—For the waiting-time to success we choose a mixture, with weights $\frac{\rho}{2+\rho}$ and $\frac{2}{2+\rho}$ respectively, of a Beta distribution on $[0, t_{\min}]$ and a Pareto distribution on $[t_{\min}, \infty)$:

$$F(t) = \frac{\rho}{2+\rho}\left(\frac{\min(t, t_{\min})}{t_{\min}}\right)^2 + I_{[t_{\min},\infty)}\frac{2}{2+\rho}F^P(t),$$

where

$$F^P(t) = 1 - \left(\frac{t}{t_{\min}}\right)^{-\rho}, \qquad \text{for } t \geq t_{\min}. \tag{43}$$

The only parameter of F is ρ. Its hazard rate is continuous and has the essential features of the bell shape in Figures (1) and (2). Then for $t < t_{\min}$, the density is $f(t) = \frac{1}{t_{\min}}\frac{2\rho}{2+\rho}\frac{t}{t_{\min}}$, and therefore (see the Appendix 3 for the derivations),

$$h(t) = \begin{cases} \frac{1}{t_{\min}}\frac{2\rho t_{\min}^{-1} t}{2+\rho-\rho t_{\min}^{-2} t^2} & \text{for } t < t_{\min} \\ \frac{\rho}{t} & \text{for } t \geq t_{\min}. \end{cases} \tag{44}$$

This distribution has two parameters: $t_{\min}$ (the age at which the success hazard peaks), and ρ (related to the mass of successes to the left of $t_{\min}$).

Fitting the termination hazard, ψ.—If wealth is distributed Pareto then we do not get the wrong sign of the Venture dummy. So we shall let wealth π be distributed as the Pareto distribution

$$G(\pi) = \Pr(\tilde{\pi} \leq \pi) = 1 - \left(\frac{\pi}{\pi_0}\right)^{-\lambda}$$

for $\pi \geq \pi_0$. To begin, let us *ignore* the lower bound on π in the derivations that follow, assuming that the parameter values are such that $\pi_{\min} \geq \pi_0$. Later we shall check that it is not violated. Then since $\gamma(\pi) = \frac{\lambda}{\pi}$, we have $\gamma\left(\frac{1+rW}{h(t)}\right) = \frac{\lambda h(t)}{1+rW}$, and we see that where ψ^{VC} and ψ^S are both positive, they are equal:

$$\begin{aligned} \psi^{VC}(t) &= (1+rW)\gamma\left(\frac{1+rW}{h(t)}\right)\left(\frac{-h'(t)}{[h(t)]^2}\right) \\ &= -\lambda\frac{h'(t)}{h(t)}. \\ &= \psi^S(t) \end{aligned}$$

In the Pareto-h case, $h(t) = \frac{\rho}{t}$, $h'(t) = -\frac{\rho}{t^2}$ and therefore

$$\psi^{VC}(t) = \psi^{S}(t) = \frac{\lambda}{t}. \tag{45}$$

Therefore in this case the VC terminates more only in the initial years.

Solving for the marginal project, $\pi_{\min}$.—As Figure 6 shows, projects for which $\pi < \pi_{\min}$ are terminated at once, whereas for the rest, T is determined by (24), so that

$$T(\pi) = \begin{cases} 0 & \text{for } \pi < \pi_{\min} \\ \tilde{\rho}\pi & \text{for } \pi \geq \pi_{\min}. \end{cases} \tag{46}$$

Appendix 1 shows that

$$\pi_{\min} > \frac{t_0}{\tilde{\rho}}. \tag{47}$$

Therefore terminations will not start until some time *after* t_0. In other words, after the initial burst of terminations at $T = 0$ we should, for a while, see successes only, and only later should terminations begin.

The distribution of terminations.—Then since (36) gives us (we drop the VC superscript now)

$$\Phi(t) = \begin{cases} 0 & \text{for } t = 0 \\ 1 - \left(\frac{\pi_{\min}}{\pi_0}\right)^{-\lambda} & \text{for } t \in (0, \tilde{\rho}\pi_{\min}) \\ 1 - \left(\frac{1}{\tilde{\rho}\pi_0} t\right)^{-\lambda} & \text{for } t \geq \tilde{\rho}\pi_{\min}. \end{cases}$$

where

$$\tilde{\rho} = \frac{\rho}{1 + rW}$$

After a spike at $t = 0$, the hazard of this distribution is zero until the point $\tilde{\rho}\pi_{\min}$ and after that it is

$$\left(\frac{t}{\tilde{\rho}}\right)^{\lambda} \left(\frac{1}{\tilde{\rho}}\right)^{-\lambda} \lambda t^{-\lambda-1} = \frac{\lambda}{t}.$$

Fitting the rates of return.—We shall fit only $R^G(t)$ as given in (30). Substituting the functional forms for h and G into (25) and (26), we compute Γ_t and then substitute that into the expression for R^G.

Fitting s^* as in (32).—We fit $s^* = \frac{1-k}{1+rW}$ to the Kaplan-Strömberg (2003, Table 2) number as described in (32) – the entrepreneur's equity share in the firm relative to that of the VC is 0.4

The rate of interest.—We fit eight things. We assume that $r = 0.127$ – the rate of return required by the CAPM model given the β of VE returns and given the S&P 500 return over the period. See the Appendix for details

Choice of units.—We maintain the assumption that $a + k = 1$, which means that our estimates are in units of total marginal costs per year. In our sample these turn out to be about \$3.5 million on average. The implications homogeneous of degree zero in the vector (a, k, π_0, C, W).

The first-period investment.—The VC payment, using (17), is

$$p = \left[1 - \frac{(1-k)}{1+rW}\right] C = \frac{k+rW}{1+rW} C.$$

But this implies that

$$\frac{\text{The VC's first-period investment}}{\text{investment in each subsequent period}} = \frac{k+p^*}{k} = 1 + \frac{k+rW}{1+rW}.$$

The parameters are ρ, $t_{\min}$, π_0, λ, k and C. The estimation algorithm is described in Appendix 1.

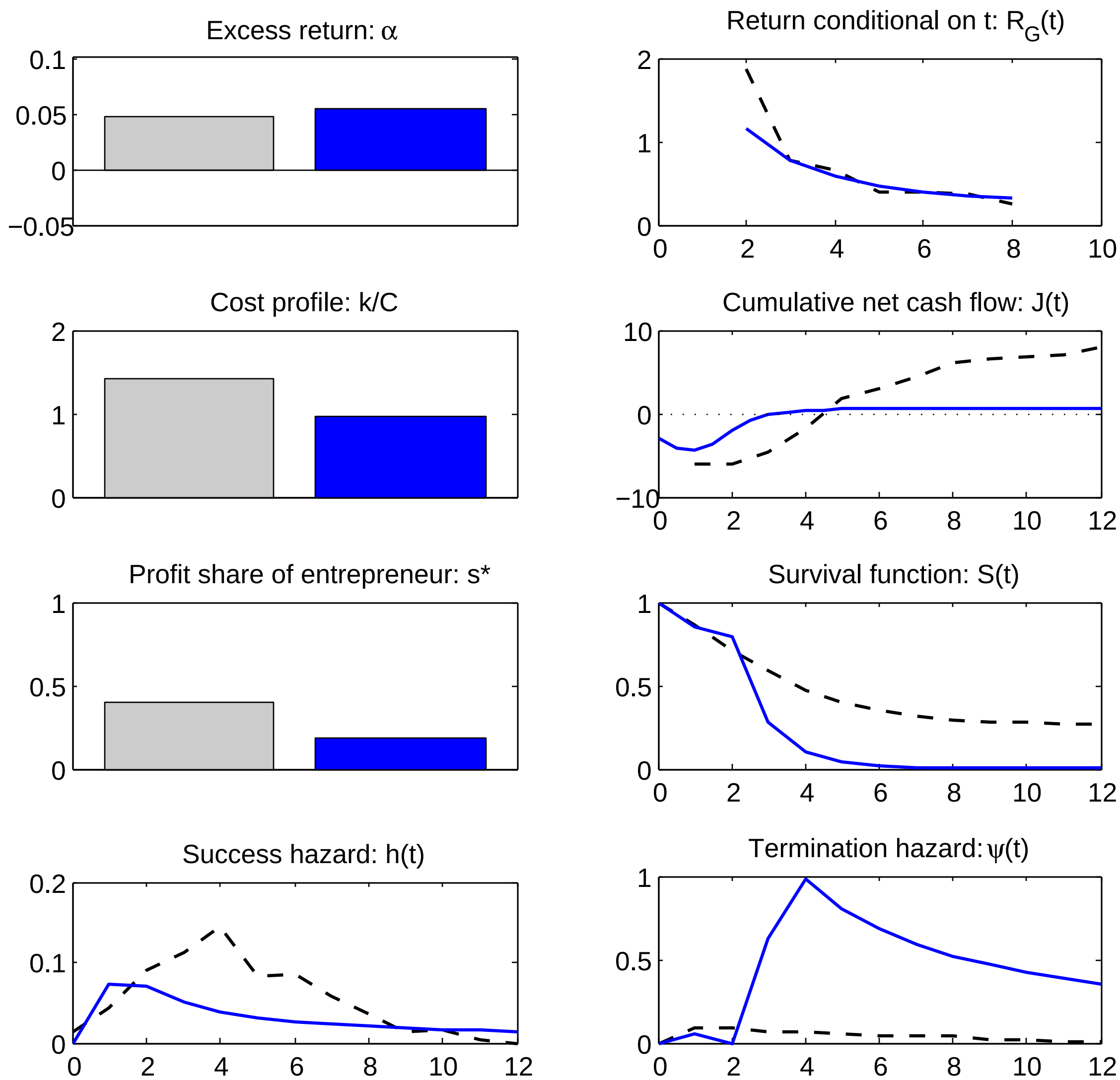
5.0.10 Estimates

The estimation procedure and the data sources are described in the Appendix. The estimates are reported together in Table 1a, and some statistics of interest are reported in Table 1B. The fit is described Visually in Figure titled "The Fit of the Pareto Model." Let us comment on each of the eight panels in the Figure on p. 36A titled "Fit of the Pareto Model." (more detail is in Appendix 1).

1. *Panel* 1.—(See Sec. 4.2, part 1). The first panel shows how well the model fits the alpha of 4.5 percent (data are the lighter, gray bar) that VE funds paid over the period 1981-95. The model (the darker, blue bar) slightly overpredicts this quantity.

2. *Panel* 2.—(See Sec. 4.2, pt. 2). The vertical axis measures the geometric rate of return in hundreds of percentage points. The horizontal axis measures the firm's age as of the date of the first VC investment. The dashed line is the return by company reported in Table A3. The highest returns are on projects that succeed early. The model (solid line) underpredicts the returns in early yeas. This is probably because VentureXpert sometimes misses investments thereby underestimating the costs. In any event, as it stands, the returns implied by the individual-firm data are higher than the VE returns that the first panel reports, and the model simply cannot fit both.

3. *Panel* 3.—The average investment, k, is about 3.5 million per year, whereas the average C is \$2.6m. The latter may seem large but the companies that a VC includes in his portfolio have often received "seed" rounds of financing before the first "real" round. The gray bar is their ratio $\frac{k}{C} = 1.35$. The model pushes for a larger C because this helps it fit more easily the cumulative cash flows and perhaps other things too.



$\alpha = 0.0476, \beta = 1.800, r_f = 0.0257, r_m = 0.0819, r = 0.1268, t_{min} = 1.8$

$\rho = 0.173, \lambda = 4.427, W = 0.475, k = 0.809, \pi_0 = 19.410, C = 0.845$

$\pi_{min} = 19.666, s^* = 0.180, \hat{p} = 0.693, k/C = 0.957, W/C_{PV} : 0.054$

Fit of the Pareto model

4. *Panel* 4.—(See Sec. 4.4). Predicted dollar flows are $\frac{1}{k}3.5$ millions of dollars. We plot medians here because for the data they are not affected much by the outliers.

5. *Panel* 5.—(See Sec. 4.5) The model underpredicts s^* because it needs a large k

6. *Panels* 6, 7 *and* 8.—(Sec. 4.7) The dashed lines in these three panels represent the data from the last three columns of Appendix Table A1 for the ratio #left/1355 (Panel 6), for $\hat{h}$ (Panel 7) and for $\hat{\psi}$ (Panel 8). The model underpredicts survival (Panel 4). Too many exits happen beyond year 3. The reason is best seen from (45) which gives the hazard at λ/t beyond a certain point. Now λ cannot get too close to unity, because then $E(\pi) = \frac{\pi_0}{\lambda - 1}$ gets too large and so do the returns. Note, too, that the date-zero termination hazard is undefined because a mass of $1 - \left(\frac{\pi_{\min}}{\pi_0}\right)^{-\lambda}$ firms are terminated immediately. We spread this mass evenly over the first year of the firm's life.

5.0.11 Discussion of the empirical results

We do not explain the heterogeneity of the returns on the venture funds. In our model, venture capital is homogeneous, though differences could easily be introduced. For instance, an experienced VC would have better signals about a project's likely success, which would lead to a more favorable distribution of waiting times, F, and of payoffs, G. Panel 1 of the Figure shows that returns drop off quickly as the waiting time increases so that an ability to bring successes forward seems to have a very high return. If VCs were different in quality because they could see a better prior signal, then the high-quality VCs would have a W higher than other VCs. The high-W VCs would then be more selective, having higher $\pi_{\min}$'s and lower $T(\pi)$s. Therefore ex post project qualities would be positively correlated with the qualities of the VCs that backed them. This would be consistent with Sørensen (2006) who finds that the bulk of the positive association bnetween VC quality and project quality is due not to direct VC influence on the payoff but to sorting.

The implied selection of entrepreneurs into venture-backed and solo projects.— At the parameter estimates, the counterpart of Figure 3 is shown in Figure 8. At \$13 million, w^* is quite large. This estimate comes about because of the number for C that we used – namely \$2.6 million. This exercise should be qualified by noting that the types of solo entrepreneurs whose projects may ever get on VCs' radar screens are not typical of all projects run by entrepreneurs; they are high-tech and require more funding than the average small business.

The implied excess return to solo entrepreneurs.—Figure 9 shows the excess return that solo entrepreneurs earn. The relevant region is to the right of w^*. we see that wealthy entrepreneurs earn a higher rate of return on their projects than poor entrepreneurs, but slightly less than α, i.e., slightly less than the VC. Thus VCs are in the right tail of the distribution, at the very top of it.

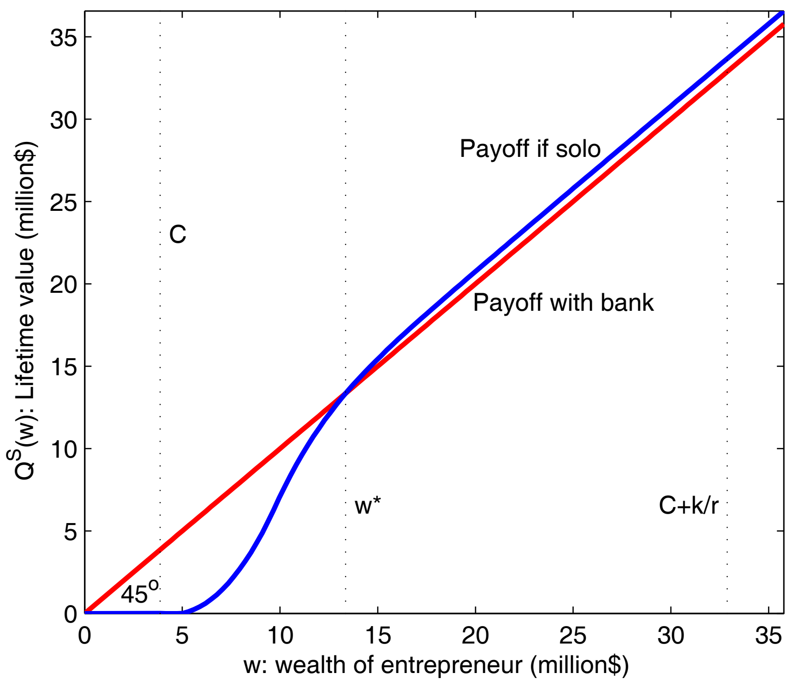


Figure 8: Estimated version of Figure 3

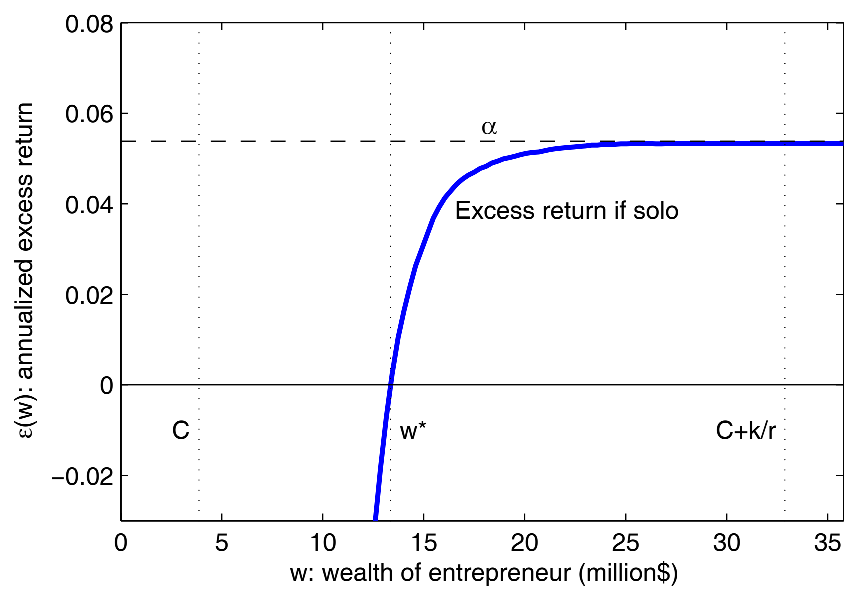


Figure 9: The excess return of solo entrepreneurs

The estimated excess return in Figure 9 is based on the assumption that all entrepreneurs – venture backed or solo – face the same costs and expected benefits, and on a revealed preference argument that we then use to extrapolate solo returns from the data on venture-backed firms *only*. In our model a solo entrepreneur expects to earn at least the market rate, and therefore she cannot have a negative alpha. There is evidence of negative alphas for some entrepreneurs (Moskowitz and Vissing-Jorgensen 2002) and independent inventors (Astebro 2003), and further work should reveal what may explain such patterns. Recently, however, Hopenhayn and Vereshchagina (2004) and Miao and Wang (2005) have argued that when one factors in the option value of the project, the alphas are nonnegative.

6 Conclusion

We estimated a model of the market for venture capital in which VCs were scarce relative to the number of potential projects. The estimates impliy a high equilibrium return on VC capital which makes the VCs impatient to start new fund and to terminate existing non-performing projects. This leads to a selection effect that gives rise to a tendency for venture-backed companies to reach IPOs earlier, and to be worth more at IPO than other start-ups..

We used the estimated model to infer the rate of return on venture capital and on entrepreneurship, the latter rising with the entrepreneur's wealth. The VC earns a higher return than even the wealthiest entrepreneur, but not by much.

The equilibrium turned out to be socially optimal for a range of parameter values.

References

[1] Astebro, Thomas. "The Return to Independent Invention: Evidence of Unrealistic Optimism, Risk Seeking or Skewness Loving?" *Economic Journal* 113, no. 484 (January 2003): 226-239.

[2] Basaluzzo, Gabriel "Entrepreneurial Teams in Financially Constrained Economies." University of Pennsylvania, November 2004.

[3] Ber, Hedva and Yafeh, Yishay. "Can Venture Capital Funds Pick Winners? Evidence from Pre-IPO Survival Rates and Post-IPO Performance." CEPR Discussion Paper No. 4672, October 2004.

[4] Bergemann, Dirk and Ulrich Hege. "VC Financing, Moral Hazard, and Learning." *Journal of Banking & Finance* 22 (1998) 703-735.

[5] ___________. "The Financing of Innovation: Learning and Stopping." *RAND Journal of Economics* 36, no. 4 (Winter 2005): 719-752.

[6] Buera, Francisco. "A Dynamic Model of Entrepreneurship with Borrowing Constraints: Theory and Evidence". Northwestern University, February 2004.

[7] Cochrane, John. "The Risk and Return of Venture Capital." *Journal of Financial Economics* 75 (January 2005): 3-52.

[8] Goldfarb, Brent, David Kirsch, and David Miller. "Was there too little entry during the Dot Com Era?" April 2006.

[9] Gompers, Paul. "Optimal Investment, Monitoring, and the Staging of Venture Capital." *Journal of Finance* 50 (1995): 1461-89.

[10] ______, and Joshua Lerner. "Money Chasing Deals? The impact of Fund Inflows on Private Equity Valuations." *Journal of Financial Economics* 55 (2000): 281-325.

[11] ______. "Risk and Reward in Private Equity Investments: The Challenge of Performance Assessment." *Journal of Private Equity* (Winter 1997): 5-12.

[12] Guler, Isin. "An Empirical Examination of Aggregate Patterns in U.S. Venture-Capital Decision Making." Mack Center, Wharton 2003.

[13] Hellmann, T., and M. Puri. "The Interaction between Product Market and Financing Strategy: The Role of Venture Capital," *Review of Financial Studies* (2000) 13, 959-984.

[14] Hochberg, Yael. "Venture Capital and Corporate Governance in the Newly Public Firm." Cornell University, September 2004.

[15] ______, Alexander Ljungqvist, and Yang Lu. "Whom You Know Matters: Venture Capital Networks and Investment Performance." *Journal of Finance*, forthcoming.

[16] Holmes, Thomas and James Schmitz. "A Theory of Entrepreneurship and Its Application to the Study of Business Transfers." *Journal of Political Economy* 98, No. 2 (Apr., 1990), pp. 265-294.

[17] Hopenhayn, Hugo and Galina Vereshchagina. "Risk Taking by Entrepreneurs." 2004.

[18] Inderst, Roman and Holger Muller. "The Effect of Capital-Market Characteristics on the Value of Start-up Firms." *Journal of Financial Economics* 72 (2004): 319–356.

[19] Jones, Charles M. and Matthew Rhodes-Kropf. "The Price of Diversifiable Risk in Venture Capital and Private Equity." GSB, Columbia University, July 2004.

[20] Kaplan, Steven and Antoinette Schoar. "Private Equity Performance: Returns, Persistence and Capital Flows." *Journal of Finance,* 2005.

[21] Kaplan, Steven and Per Strömberg. "Financial Theory Meets the Real World: An Empirical Analysis of Venture Capital Contracts." *Review of Economic Studies* 70(2) no. 243 (April 2003): 281-316.

[22] Kaplan, Steven N., Berk A. Sensoy, and Per Strömberg. "How well do venture capital databases reflect actual investments?" University of Chicago, September 2002.

[23] Lerner, Joshua. "The syndication of venture capital investments." *Financial Management* 23 (Autumn 1994): 16-27.

[24] ___________. "Comment on Bergemann and Hege." *Journal of Banking & Finance* 22 (1998) 736-40.

[25] ___________, Antoinette Schoar, and Wan Wong. "Smart Institutions, Foolish Choices?: The Limited Partner Performance Puzzle" *Journal of Finance*, forthcoming.

[26] Ljungqvist, Alexander and Matthew Richardson. "The Cash Flow, Return and Risk Characteristics of Private Equity." NBER Working Paper No. 9454, January 2003.

[27] Miao, Jianjun and Neng Wang. "Learning, Investment and Entrepreneurial Survival." October 2005.

[28] Michelacci, Claudio and Javier Suarez. "Business Creation and the Stock Market." *Review of Economic Studies* 71, no. 2 (2004): 459-481.

[29] Ruhnka, John C. & Feldman, Howard D. & Dean, Thomas J., 1992. "The "living dead" phenomenon in venture capital investments," *Journal of Business Venturing*, Elsevier, vol. 7(2), pages 137-155.

[30] Sørensen Morten. "How Smart is Smart Money? A Two-Sided Matching Model of Venture Capital." U. Chicago, January 2006.

[31] Statman, Meir, and James F. Sepe, 1989. "Project termination announcements and the market value of the firm." *Financial Management* 18 (1989): 74-81.

7 Appendix

7.1 Data and the estimation algorithm

In general, we will denote the quantities implied by the model by plain letters (x, e.g.), whereas the moments from the data are denoted by the same letter carrying a hat ($\hat{x}$).

The data on α and r.—The value $\hat{\alpha} = 0.0468$ that we use is inferred from the quarterly value $\alpha_q = 0.0117$ that Jones and Rhodes-Kropf (2004) report on the bottom of panel A in table 2. In order to compute the discount rate $r = 0.127$ that we use in the computations, we create the typical return for an investment with the characteristics of VC projects following the capital asset pricing model (CAPM):

$$r = r_f + \beta_{VC}(r_m - r_f)$$

We compute r_f as the mean of the return on 3-month treasury bills from 1980 to 1999, which is the sample period for Jones and Rhodes-Kropf (2004)'s data. Similarly, we compute r_m as the mean of the return on a value-weighted portfolio of stocks listed on the NYSE, AMEX and NASDAQ provided by the Center for Research in Security Prices (CRSP) over the same period. The value $\beta_{VC} = 1.80$ is calculated as the sum of the 5 quarterly β-coefficients that Jones and Rhodes-Kropf (2004) provide in table 1, panel B.

The data on τ and T.—These data are from the VentureExpert database provided by Venture Economics, and are described in detail by Guler (2003). The following table summarizes the data on successes and terminations. Age is measured as the number of periods since the date of first investment.

age	ipo	acq	term	#evnts	#left	$\hat{h}$	$\hat{\psi}$
0	12	8	0	20	1355	0.01	0.00
1	39	19	119	177	1335	0.04	0.09
2	54	49	103	206	1158	0.09	0.09
3	65	42	61	168	952	0.11	0.06
4	67	47	50	164	784	0.15	0.06
5	27	24	36	87	620	0.08	0.06
6	22	23	20	65	533	0.08	0.04
7	16	11	19	46	468	0.06	0.04
8	5	10	17	32	422	0.04	0.04
9	0	5	6	11	390	0.01	0.02
10	2	4	6	12	379	0.02	0.02
11	0	1	1	2	367	0.00	0.00
12	0	0	1	1	365	0.00	0.00

Table A1

The last three columns are plotted as the dashed lines in panels six (there normalized by dividing by 1355), seven and eight of Figure **??**.

The data on k/C.—As a typical value for C, we take $\hat{C} = \$2.6$m. This number is reported by Kaplan, Sensoy and Strömberg (2002) in table 2 as the book value of a VC portfolio company at the business plan in the median. In order to get an estimate $\hat{k}$ for the typical monetary investment k that a portfolio company requires each year, we employ the following data from Guler (2003, Table 6, column 2):

Investment round	1	2	3	4	5	6	7	8	9	10	11	12
Amount ($ millions)	6.5	5.3	5.5	7	8.4	6.4	5.9	8.2	3.4	8.1	3.4	3.4

Table A2

Because the model has k paid per unit of time and C at the outset, we need to convert these into spending per year. By comparing the speed of terminations we arrived at the conversion factor for converting rounds into flows. If I_j is the average amount invested in round j, we convert this into a flow $I_t = \theta_t I_t$, where[14]

$$\theta_t \equiv \frac{1}{1.25} - \frac{1}{5}\left(\frac{1}{1.25} - \frac{1}{1.5}\right)t.$$

As an estimate for the typical yearly investment flow once the initial investment round is over, we obtain $k = \frac{1}{T-1}\sum_{t=2}^{T} \theta_t I_t$.

The data on $R_G(t)$.—These also come from Guler (2003)'s data. They are annualized log returns for the projects in Guler's sample and we reproduce them in Table A3:

Year	1-2y	2-3y	3-4y	4-5y	5-6y	6-7y	7-8y
Geometric return	1.861	0.785	0.654	0.401	0.395	0.370	0.258

Table A3

The data are also plotted as the dashed line in the second panel of Figure **??**.

The data on $J(t)$.—To estimate the cumulative cash-flow $J(t)$ of a typical project, we take the average cumulative cash-flow in Guler's data up to year t. Guler's data for π are the market valuations for the respective company at IPO or at acquisition. In line with the value $s^* = 0.40$ that we use (see below), we assume that 60% of these numbers goes to VCs. Furthermore, for the positive cash-flows in the formula for $J'(t)$ in section 4.0.5 we calculate *conditional medians at t* instead of *conditional means* — we do this because the mean is very sensitive to the extreme outliers that are present in Guler's data. We reproduce the series we obtain for $\hat{J}(t)$ in table A4:

[14]Between the first and the sixth round,the termination hazard falls from 0.12 to 0.08. On the other hand, between year 1 and year 6, the termination hazard falls from 0.09 to 0.04. Thus the ratio of the two hazards rises from $\frac{12}{9} = 1.25$ to $\frac{8}{4} = \frac{1}{2}$. As a rough calculation, then, initially, rounds are once every 1.25 years, and by year 6, they are once every 1.5 years.

Year	1	2	3	4	5	6	7	8	9	10	11	12	
$\hat{J}_t$	-6.1	-6.0	-4.6	-1.6	1.7	2.9	4.3	6.2	6.6	6.9	7.1	7.9	*Table A4*

The data are also plotted as the dashed line in the fourth panel of Figure **??**.

The data on s^*.—The reference value $\hat{s}^* = 0.40$ is taken from Kaplan and Strömberg (2002), table 2.

Computations.—Here is how we calculated the hazards:

1. Column 6, "# left" is the empirical counterpart of $(1 - F[t])(1 - \Phi[t])$, i.e., of $S(t)$ in (42).

2. Column 7, "$\hat{h}$" is the ratio $(\texttt{ipo} + \texttt{acq})/(\texttt{\#left})$. So, e.g., the value of this ratio at age 1 is $\frac{39+19}{1335} = 0.043$, its value at age 2 is $\frac{54+49}{1158} = 0.089$, and so on. We now show that this is the empirical counterpart of $h(t)$. The sum of the columns $(\texttt{ipo} + \texttt{acq})$ we interpret as the number of successes at date t among firms for whom $T > t$. The probability of $\tau = t$ *and* its surviving beyond t is $f(t)(1 - \Phi[t])$. Therefore we equate these two concepts:

$$(\texttt{ipo} + \texttt{acq}) = f(t)(1 - \Phi[t]).$$

   Therefore as calculated in Column 7, the "success hazard" is

$$\frac{\texttt{ipo + acq}}{\texttt{\# left}} = \frac{f(t)(1 - \Phi[t])}{(1 - F[t])(1 - \Phi[t])} = \frac{f(t)}{1 - F(t)} = h(t).$$

3. Column 8, "$\hat{\psi}$" is the ratio $(\texttt{term})/(\texttt{\#left})$. This is the empirical counterpart of $\psi(t)$; to see why, note that term is the number terminated at date t among firms for whom $T \geq t$ and $\tau \geq t$. The probability of $T = t$ *and* $\tau \geq t$ is $\Phi'(t)(1 - F[t])$. Therefore we equate the two concepts: (term) $= \Phi'(t)(1 - F[t])$ whereupon, as calculated in Column 8, the "termination hazard" is

$$\frac{\texttt{term}}{\texttt{\# left}} = \frac{\Phi'(t)(1 - F[t])}{(1 - F[t])(1 - \Phi[t])} = \frac{\Phi'(t)}{1 - \Phi[t]} = \psi(t).$$

Solving for the marginal project, $\pi_{\min}$.—As Figure 6 shows, projects for which $\pi < \pi_{\min}$ are terminated at once, whereas for the rest, T is determined by (24), so that (46) holds. In (1), an integration by parts leads to

$$\begin{aligned}\int_0^T e^{-rt} f(t)\,dt &= -\,e^{-rt}(1 - F[t])\Big|_0^T + r\int_0^T e^{-rt}(1 - F[t])\,dt \\ &= 1 - e^{-rT}(1 - F[T]) + r\int_0^T e^{-rt}\frac{f(t)}{h(t)}dt,\end{aligned}$$

so that when substituted into (1) yields

$$V(\pi) \equiv W + \max_{T} \int_0^T \left(\pi - \frac{1+rW}{h(t)}\right) e^{-rt} f(t)\, dt$$

Substituting the Pareto form for h, the social surplus that project π delivers is

$$\begin{aligned} V(\pi) - W &= -\int_0^{t_0} e^{-rt} dt + \int_{t_0}^{\infty} e^{-rt} \max\left(0, \pi - \frac{t}{\tilde{\rho}}\right) \rho t_0^{\rho} t^{-1-\rho} dt \\ &= -\frac{(1-e^{-rt_0})}{r} + \rho t_0^{\rho} \int_{t_0}^{\infty} e^{-rt} \left(\pi - \frac{t}{\tilde{\rho}}\right) t^{-1-\rho} dt. \qquad (48) \end{aligned}$$

The marginal project $\pi_{\min}$ therefore solves $V(\pi) = W$. Substituting $\pi = t/\tilde{\rho}$ on the RHS of (48), we see that $V(t_0/\tilde{\rho}) < W$; since V is increasing in π, (47) follows.

The geometric rate of return.—The mean geometric return of successes for a fixed success time t can be obtained in closed form if π is distributed Pareto (use equations (29), (30) and (26) to obtain this):

$$R(t,\pi) = \frac{\ln(\pi)}{t} + \frac{\ln(1-s^*) - \ln\left[\frac{k}{r}(1-e^{-rt}) + C - p^*\right]}{t}$$

To obtain a closer analog to the data, however, we average out over a whole year using the density of successes:

$$R_G^t = \int_{t-1}^{t} R_G(t) f(t) dt \simeq \sum_{i=1}^{N} R_G(t+i/N) f(t+i/N)$$

In the program, we use N=10 for the approximation of the integral.

VC Value and termination policy.—The value to the VC of an *accepted* project with payoff π and terminated at T is

$$\begin{aligned} V_a(T|\pi, W) &= \int_{t_0}^{T} \left[(\pi + W)e^{-rt} - \int_0^t e^{-rs} ds\right] f(t) dt \\ &\quad + (1-F(T))(e^{-rT}W - \int_0^T e^{-rt} dt) \\ &= \int_{t_0}^{T} \left[(\pi + W)e^{-rt} - (1-e^{-rt})/r\right] f(t) dt \\ &\quad + (1-F(T))(e^{-rT}W - (1-e^{-rT})/r) \\ &= \int_{t_0}^{T} (\pi + W + \frac{1}{r}) e^{-rt} f(t) dt + (1-F(T)) e^{-rT}(W + 1/r) - \frac{1}{r} \end{aligned}$$

The optimal stopping time solves

$$\frac{dV_a(T|\pi,W)}{dT}=0=e^{-rT}f(T)\left(\pi-\frac{1+rW}{h\,(T)}\right),\text{ if }T>t_0.$$

For the Pareto distribution $T=\pi\rho/(1+rW)$, and so

$$V_a(\pi,W)=\rho t_0^{\rho}\int_{t_0}^{T}(\pi+W+\frac{1}{r})\frac{e^{-rt}}{t^{\rho+1}}dt+(1-F(T))e^{-rT}(W+\frac{1}{r})-\frac{1}{r}$$

Estimation Strategy.—The set of unknown parameters is $\{\rho,\pi_0,\lambda,C,k\}$. To estimate them, we fix $t_0=1.8$ and $r=0.127$.[15] We exploit the monotonous relationship between C and the model outcome W apparent in equation (13) for our estimation algorithm: Instead of solving the model for a given C, we choose $W\geq 0$ and back out the C that is compatible with it given the parameters (ρ,π_0,λ,k).

To compute the loss function RSS for a set of parameters (ρ,π_0,λ,k,W) given in (49), we follow the following steps:

1. Using $\{\rho,\pi_0,\lambda,W\}$ compute $\pi_{\min}$ from $V_a(\pi_{\min};W)=W$.
2. Using $\{\rho,\pi_0,\lambda,W,\pi_{\min}\}$ compute C according to

$$\begin{aligned}C=&\int_{\pi_{\min}}^{\infty}\Big[\int_{t_0}^{T}e^{-rt}(\pi+W+\frac{1}{r})f(t)dt+(1-F(T))e^{-rT}(W+\frac{1}{r})\Big]g(\pi)d\pi\\&-(1-G(\pi_{\min}))(W+\frac{1}{r}).\end{aligned}$$

3. Using $\{\rho,\pi_0,\lambda,W,k,\pi_{\min}\}$ compute p^* via

$$\begin{aligned}p^*=&(1-k)\int_{\pi_{\min}}^{\infty}\Big[\int_{t_0}^{T}e^{-rt}\Big(\frac{\pi}{1+rW}+\frac{1}{r}\Big)f(t)dt+(1-F(T))\frac{e^{-rT}}{r})\Big]g(\pi)d\pi\\&-(1-k)(1-G(\pi_{\min}))/r.\end{aligned}$$

4. Compute the present expected value $C_{PV}=C_{PV}^{M}+C_{PV}^{C}$ of all costs for the VC: Clearly, the part coming from marginal cost must be $C_{PV}^{M}=k/r$, since the VC will never be idle and hence always pay a flow cost k. The part C_{PV}^{C} stemming

[15] The choice of r has been described above. We choose a low value for t_0 since terminations cannot happen before t_0; hence if we choose a value above 2, then the model will definitelypredict zero terminations for projects of age 1 to 2. On the other hand, t_0 marks the maximimum of the success hazard, which occurs rather late in the data, so we move t_0 rather close to 2. In subsequent versions of the paper, we also want to maximize with respect to this parameter within reasonable bounds.

from the repeated payment of the fixed cost C can be obtained in a recursive fashion:

$$C_{PV}^{C} = C + E[e^{-rT} C_{PV}^{C}],$$

where $T = \min\{\tau, T^*(\pi)\}$ is the time of success or termination of the project, whichever comes first. From this, we obtain

$$C_{PV}^{C} = \frac{C}{1 - E[e^{-rT}]}, \qquad E[e^{-rT}] = \int_0^{\pi} e^{-r\min\{\tau, T^*(\pi)\}} g(\pi) d\pi.$$

$E[e^{-rT}]$ can be computed in a manner quite similar to C, so it is not explicitly given here.

5. Now, we can obtain the α implied by the model: $\alpha = W/C_{PV}$, as given in equation (27).

6. The proportion of fixed cost to flow investment is k/C.

7. Compute $R_G(t)$ for $t = 2, 3, \ldots, 12$ as described before.

8. Compute $J(t)$ for $t = 1, 2, \ldots, 12$ by numerically integrating as follows:

$$J(t) = -\hat{p} + \int_0^t J'(s) ds \simeq -\hat{p} + \sum_{i=1}^{t/\Delta t} J'\big[(i-1)\Delta t\big]\Delta t,$$

where we choose $\Delta t = 0.01$ and where $J'(t)$ is given in section 4.0.5 and can be evaluated in closed form.

9. The predicted share of the entrepreneur in the profit π is

$$s^* = \frac{1-k}{1+rW}.$$

10. Using $\{\rho, t_0\}$ compute $h_t, t = 0, 1, .., 12$ according to the formula given in section 7.3. For the first two periods, we calculate the hazard as follows to obtain a higher degree of precision:

$$h_t = \frac{F(t) - F(t-1)}{F(t-1)},$$

where $F(t)$ is the distribution function of successes obtained from the model.

11. Using $\{\rho, t_0, \lambda, W, \pi_{\min}\}$ compute the survival function S_t, $t = 0, 1, .., 12$, as given above.

12. Compute the criterion function:

$$\begin{aligned}\mathrm{RSS} = & \, w_1[\ln\hat{\alpha} - \ln\alpha]^2 + w_2\big[\ln(\hat{k}/\hat{C}) - \ln(k/C)\big]^2 + \\ & + w_3\sum_{i=2}^{N}[\hat{R}_G(t) - R_G(t)]^2 + w_4\sum_{t=0}^{12}[\hat{J}_t - J_t]^2 + w_5[\ln\hat{s^*} - \ln s^*]^2 + \\ & + w_6[\ln\hat{\rho} - \ln\rho]^2 + w_7\left([\ln\hat{S}_1 - \ln S_1]^2 + [\ln\hat{S}_5 - \ln S_5]^2\right)\end{aligned} \tag{49}$$

We chose penalties on logarithms for some moments since a simple quadratic scheme did not penalize absurd behavior sufficiently; for example, when the ratio k/C goes from 1/10 to 1/100, this is a big qualitative change in our eyes but would not be penalized much by a purely quadratic scheme.

Since t_{min} is fixed, the success hazard h_t is governed solely by the parameter ρ, hence we simplify the loss function by bringing ρ close to $\hat{\rho} = 0.3$, which yields a good fit of the data. It turned out that terminations could be more efficiently penalized via the survival function S_t, so we chose to penalize survival at two critical points: After year one, the initial burst of terminations is compared to the data; year five provides another data point that checks the termination subsequent terminations predicted by the model.[16]

For the weighting scheme, we chose $w = (10, 1, 1, 0.05, 1, 1, 1)$. The high penalty on α reflects both the importance of this parameter and its low scale. The low penalty on the J-curve, in turn, was chosen because of the large scale of $J(t)$ and the lesser importance we attach to it in the estimation.

This weighting scheme yielded the estimates reported in Table 2A.

Estimates of R, S, and h are plotted in the eight-panel Figure.

Details on how C and p^* were computed.—The formula for C is

$$\begin{aligned}C &= \int_{\pi_{\min}}^{\infty} V_a(\pi, W)g(\pi)d\pi - (1 - G(\pi_{\min}))W \\ &= \int_{\pi_{\min}}^{\infty}\left[\int_{t_0}^{T(\pi)}\Big(\pi + W + \frac{1}{r}\Big)e^{-rt}f(t)dt + (1 - F(T(\pi)))e^{-rT(\pi)}\Big(W + \frac{1}{r}\Big) - \frac{1}{r}\right]g(\pi)d\pi \\ &\quad - (1 - G(\pi_{\min}))W \\ &= S_2 + \Big(W + \frac{1}{r}\Big)(S_1 + S_3 - 1 + G(\pi_{\min})),\end{aligned}$$

[16] Note that the success and the termination hazard together determine the survival function.

where

$$S_1 = \int_{\pi_{\min}}^{\infty} \Big[\underbrace{\int_{t_0}^{T(\pi)} e^{-rt} f(t) dt}_{\equiv \gamma(\pi)} \Big] g(\pi) d\pi$$

$$S_2 = \int_{\pi_{\min}}^{\infty} \pi \int_{t_0}^{T(\pi)} e^{-rt} f(t) dt \; g(\pi) d\pi$$

$$S_3 = \int_{\pi_{\min}}^{\infty} (1 - F(T(\pi))) e^{-rT(\pi)} g(\pi) d\pi$$

Then, similarly,

$$\begin{aligned} p^* &= (1-k)\Big[S_2/(1+rW) + \frac{1}{r}(S_1 + S_3 - 1 + G(\pi_{\min}))\Big] \\ &= (1-k)\Big[S_2 + \Big(W + \frac{1}{r}\Big)\Big(S_1 + S_3 - 1 + G(\pi_{\min})\Big)\Big]/(1+rW) \\ &= (1-k)C/(1+rW) \end{aligned} \tag{50}$$

Note that the integrals S_1, S_2 and S_3 do not have closed-form solutions. The example S_1 shows how we evaluate these integrals: The inner integral $\gamma(\pi)$ involves integration against a function of type e^{-rt}, thus –after a simple change of variable– *Laguerre Quadrature* can be used to calculate $\gamma(\pi)$ to a very high degree of precision for fixed π. To calculate the outer integral, we make use of the fact that the function $\gamma(\pi)$ converges to a linear asymptote for large π; an integral of a linear function against $g(\pi)$, however, can be evaluated in closed form. We determine the point $\bar{\pi}$ where $\gamma(\pi)$ satisfies some convergence criterion to the asymptote and evaluate $\int_{\bar{\pi}}^{\infty} \gamma(\pi) g(\pi) d\pi$ using this linear approximation. The finite part $\int_{\pi_{min}}^{\bar{\pi}} \gamma(\pi) g(\pi) d\pi$ is then evaluated using *Legendre Quadrature*.

We use 30 nodes to calculate the quadrature approximations for all integrals. Checks with the *Matlab*-built-in *adaptive Simpson Quadrature* (for the single integrals) and Monte-Carlo integration (for the double integrals) showed that the error in the calculations using these methods are of negligible order. Laguerre and Legendre are computationally inexpensive, so we preferred to use them in the minimization runs.

We use a *Matlab*-built-in line-search method to minimize RSS with respect to the parameters of the model. The minimization process proved to be rather robust; the algorithm converged to the same solution for almost any randomly chosen starting point.

Using numerical gradients and standard errors of the moments we try to match, it is possible to obtain (approximate) standard errors for our estimators. This is an important next step in this project.

7.2 Proof of Lemma 4

(i) Notice that, given that a project is supported up to time T,

$$\begin{aligned}
V(\pi) - V^{VC}(\pi) &= \int_0^{T(\pi)} \left(\pi - \frac{1}{h(t)} - \left[(1-s)\pi - \frac{k}{h(t)}\right]\right) e^{-rt} f(t)\, dt \\
&= \int_0^{T(\pi)} \left(\pi - \frac{1}{h(t)} - \left[\left(1 - \frac{a}{1+rW}\pi\right) - \frac{k}{h(t)}\right]\right) e^{-rt} f(t)\, dt \\
&= a\int_0^{T(\pi)} \left(\frac{1}{1+rW}\pi - \frac{1}{h(t)}\right) e^{-rt} f(t)\, dt \quad \text{(because } 1-k=a\text{).}
\end{aligned}$$

Multiplying through by $(1+rW)/a$,

$$\begin{aligned}
0 &= V(\pi) - V^{VC}(\pi) \quad \Longleftrightarrow \\
0 &= \int_0^{T(\pi)} \left(\pi - \frac{1+rW}{h(t)}\right) e^{-rt} f(t)\, dt \\
&= \int_0^{T(\pi)} \left(\pi - \frac{1}{h(t)}\right) e^{-rt} f(t)\, dt + W\int_0^{T(\pi)} -re^{-rt}(1-F[t])\, dt. \qquad (51)
\end{aligned}$$

Now, integrating by parts, one can rewrite the last expression in the previous equality chain as

$$\begin{aligned}
\int_0^{T(\pi)} -re^{-rt}(1-F[t])\, dt &= e^{-rt}(1-F[t])\Big|_0^{T(\pi)} + \int_0^{T(\pi)} e^{-rt} f(t)\, dt \\
&= e^{-rT(\pi)}(1-F[T(\pi)]) - 1 + \int_0^{T(\pi)} e^{-rt} f(t)\, dt \qquad (52)
\end{aligned}$$

Substituting from (52) into (51) we see, that (51) reads

$$\begin{aligned}
0 &= \int_0^{T(\pi)} \left(\pi - \frac{1}{h(t)}\right) e^{-rt} f(t)\, dt + W\left(e^{-rT(\pi)}(1-F[T(\pi)]) - 1 + \int_0^{T(\pi)} e^{-rt} f(t)\, dt\right) \\
&= \int_0^{T(\pi)} \left(\pi - \frac{1}{h(t)}\right) e^{-rt} f(t)\, dt + W\left(e^{-rT(\pi)}(1-F[T(\pi)]) + \int_0^{T(\pi)} e^{-rt} f(t)\, dt\right) - W \\
&= \int_0^{T(\pi)} \left(\pi + W - \frac{1}{h(t)}\right) e^{-rt} f(t)\, dt + We^{-rT(\pi)}(1-F[T(\pi)]) - W \\
&= \frac{1}{a}(V(\pi) - W).
\end{aligned}$$

Therefore (51) and (52) imply that

$$0 = V(\pi) - V^{VC}(\pi) \quad \Longleftrightarrow 0 = V(\pi) - W \qquad (53)$$

Since $V(\pi_{\min}) = W$, this implies claim (i).

(ii) If $s = a/(a+k+rW)$ then

$$\begin{aligned} & \int_0^T \left(\frac{a}{a+k+rW}\pi - \frac{a}{h(t)} \right) e^{-rt} f(t)\,dt \\ = & \frac{a}{1+rW} \int_0^T \left(\pi - \frac{1+rW}{h(t)} \right) e^{-rt} f(t)\,dt. \end{aligned}$$

But this is exactly $V(\pi) - V^{VC}(\pi)$ (see (51)) which, by claim (i) of this Lemma was shown to be zero at $\pi_{\min}$. ■

7.3 Derivation of the hazard for the estimated example

We use n as the exponent in order to clarify the algebra, and then we shall evaluate the result at $n = 2$. We write the mixing parameter as μ. Then

$$F(t) = (1-\mu)\left(\frac{\min(t, t_{\min})}{t_{\min}}\right)^n + I_{[t_{\min},\infty)}\mu F^P(t),$$

Then For $t < t_{\min}$, $f(t) = \frac{1}{t_{\min}}(1-\mu)n\left(\frac{t}{t_{\min}}\right)^{n-1}$, and therefore

$$\begin{aligned} \frac{f(t)}{1-F(t)} &= \frac{1}{t_{\min}} \frac{(1-\mu)n\left(\frac{t}{t_{\min}}\right)^{n-1}}{1-(1-\mu)\left(\frac{t}{t_{\min}}\right)^n} = \frac{1}{t_{\min}} \frac{(1-\mu)nt_{\min}^{1-n}t^{n-1}}{1-(1-\mu)t_{\min}^{-n}t^n} \\ &= \frac{(1-\mu)nt_{\min}^{1-n}t^{n-1}}{t_{\min}-(1-\mu)t_{\min}^{1-n}t^n} \end{aligned}$$

and

$$\lim_{t \nearrow 1} \frac{f(t)}{1-F(t)} = \frac{1}{t_{\min}} \frac{1-\mu}{\mu} n$$

For $t \geq t_{\min}$

$$\begin{aligned} \frac{f(t)}{1-F(t)} &= \frac{\mu\rho t_{\min}^{\rho} t^{-\rho-1}}{1-\left[(1-\mu)+\mu\left(1-\left(\frac{t}{t_{\min}}\right)^{-\rho}\right)\right]} \\ &= \frac{\mu\rho t_{\min}^{\rho} t^{-\rho-1}}{1-1+\mu-\mu+\mu\left(\frac{t}{t_{\min}}\right)^{-\rho}} \\ &= \frac{\rho t_{\min}^{\rho} t^{-\rho-1}}{\mu-\mu\left(1-\left(\frac{t}{t_{\min}}\right)^{-\rho}\right)} = \frac{\rho}{t} \end{aligned}$$

Therefore the hazards are equal at $t_{\min}$ if

$$\frac{1}{t_{\min}}\frac{1-\mu}{\mu}n=\frac{\rho}{t_{\min}},$$

i.e., if

$$\mu=\frac{1}{1+\frac{\rho}{n}}$$

After setting $n=2$, this leads to $\mu=\frac{2}{2+\rho}$, $1-\mu=\frac{\rho}{2+\rho}$, and, hence for $t<t_{\min}$,

$$\frac{f}{1-F}=\frac{1}{t_{\min}}\frac{(1-\mu)\,nt_{\min}^{1-n}t^{n-1}}{1-(1-\mu)\,t_{\min}^{-n}t^{n}}=\frac{1}{t_{\min}}\frac{\frac{\rho}{2+\rho}nt_{\min}^{-1}t}{1-\frac{\rho}{2+\rho}t_{\min}^{-2}t^{2}}=\frac{1}{t_{\min}}\frac{2\rho t_{\min}^{-1}t}{2+\rho-\rho t_{\min}^{-2}t^{2}}$$

which leads to to (44).

7.4 Details on $R\,(t)$ and its relation to the IRR

The *arithmetic* rate of return satisfies the equation

$$R^{A}\,(t)=\ln\int e^{R(t,\pi)}d\Gamma_{t}\,(\pi)\,.$$

Neither R^{G} nor R^{A} can be said to be correct or incorrect; each attempts to measure, in a single number, the properties of a distribution. because the function e^{R} is convex in R, Jensen's inequality implies that $R^{A}\,(t)\geq R^{G}\,(t)$, with strict inequality if Γ_{t} has positive variance. [17]

In a finite sample of projects with their π's drawn from Γ_{t}, the realized $R^{G}\,(t)$ and $R^{A}\,(t)$ would deviate from their theoretically-predicted values, but Cochrane has a fairly large sample, at least for the successes registered fairly early on in the firms' lives. Over sufficiently many projects that lasted t periods *and* succeeded, this would roughly be the realized rate of return.

The IRR on projects that succeed at age t.—Parallel to the definition of the rate of return in (29), we can define the internal rate of return on the quality-π project that matures at t; call it IRR(τ,π), as solving the equation

$$e^{\mathrm{IRR}(t,\pi)t}=\frac{(1-s^{*})}{\int_{0}^{t}ke^{-\mathrm{IRR}(t,\pi)u}du+C-p^{0}}\pi. \tag{54}$$

[17] For instance (and this is Cochrane's assumption, but it is not consistent with our model, as is evident from Figure 6), if $\Gamma_{t}\,(\pi)$ were the normal distribution, then $R\,(t,\pi)$ would be normally distributed with mean $R^{G}\,(t)$, and with a variance the we shall denote by $Var_{t}\,(R)$. Then $\int e^{R(t,\pi)}d\Gamma_{t}\,(\pi)=e^{R^{G(t)}+\frac{1}{2}Var_{t}(R)}$, and we would then have

$$R^{A}\,(t)=R^{G}\,(t)+\frac{1}{2}Var_{t}\,(R)\,.$$

Note the difference in the denominators of (29) and (54). Note that we can, once again, have an arithmetic and geometric concepts of the IRR. We shall not fit the IRR in this paper although Guler (2003) calculates it in her paper for the companies in her sample.

Proposition 4 *$R(t,\pi)$ and $IRR(t,\pi)$ are strictly increasing in π and strictly decreasing in t*

Proof. From (29), $\ln R(t,\pi) = \frac{1}{t}\left(\ln\left[(1-s^*)\pi\right] - \ln\left[\int_0^t ke^{-ru}du + C - p^0\right]\right)$ and the claim for R follows. Multiplying both sides of (54) by the denominator of its RHS leads to

$$\int_0^t ke^{\text{IRR}(t,\pi)(t-u)}du + e^{\text{IRR}(t,\pi)t}\left(C - p^0\right) = (1-s^*)\,\pi.$$

By Lemma 3,[$C - p^0 = \frac{kC}{1+rW} > 0$. Therefore The LHS is increasing in IRR(τ,π) and in t. The RHS is increasing in π, and so the claim for IRR follows as well. ■

Unfortunately, while for π fixed, $R(t,\pi)$ is declining in t, we cannot prove in general that $R^G(t)$ and $R^A(t)$ decline in t. The reason is the selection effect on π that Figure 6 portrays. Projects that last longer are subject to more stringent selection – the truncation point $(1+rW)/h(t)$ moves to the right as the products age. This positive selection effect may offset the fact that older projects have higher cumulative costs. In the estimated model the denominator effect easily dominates, and R^G and R^A both decline rapidly with t.

The relation between the rate of return and the IRR.—For completeness, we shall add a tangential result. In the (empirically relevant) parameter range for which the IRR exceeds the outside rate of interest, r, the IRR also exceeds the rate of return:

Lemma 7 *For each (τ,π),*

$$IRR(t,\pi) > R(t,\pi) \quad \text{if and only if} \quad IRR(t,\pi) > r.$$

Proof. Since IRR$(t,\pi) > r$, $e^{-\text{IRR}(t,\pi)u} < e^{-ru}$ for all $u \geq 0$. Then the denominator in (54) is smaller than the denominator in (29) and the claim follows. ■