PRELIMINARY DRAFT

COMMENTS WELCOME

# Learning by Investing: Evidence from Venture Capital

Morten Sorensen[†]

August, 2006

**Abstract:** Venture capital investors (VCs) can create value by actively exploring new investment opportunities to learn about their returns. In traditional financial markets, a free-rider problem reduces exploration and learning, but VCs' organizational structure may limit information spillovers and reduce this problem. I present a basic model of learning, based on the statistical Multi-Armed Bandit model. The value of an investment consists of both its immediate return and an option value of learning. The model is estimated, and it is found that VCs who explore more have higher returns.

Venture capitalists (VCs) invest in privately held entrepreneurial companies and are actively involved in monitoring and managing these companies.[1] The literature has identified a number of ways VCs add value. VCs screen for good companies, and their involvement with these companies is valuable in several ways: They are active board members (Lerner (1995)); they help bring products to market faster (Hellmann and Puri

---

[†] Assistant Professor, University of Chicago, GSB. Email: msorense@ChicagoGSB.edu. I want to thank Jean-Pierre Dube, Gunter Hitsch, Steve Kaplan, Josh Lerner, Laura Lindsey and Francisco Perez-Gonzalez for helpful comments and discussions. I am grateful to seminar participants at Arizona State University, University of Illinois–Chicago, and the Chicago GSB Faculty Lunch for helpful comments and to Susan Woodward of Sand Hill Econometrics for generous access to data.

[1] See Gompers and Lerner (1999), Gorman and Sahlman (1989), and Sahlman (1990) for descriptions of the VC industry and the role of VC investors in general.

(2000)); they replace inefficient management (Hellmann and Puri (2002)); and the reputation of an established VC may certify the value of a young entrepreneurial company (Megginson and Weiss (1991)). Kaplan and Strömberg (2004) find direct evidence of VCs' contributions from their investment analyses.

This paper proposes an additional way VCs create value. Specifically, the organizational structure of VC investors allows them to internalize benefits of learning and provides incentives for exploring new ideas and technologies, which is socially valuable. When investing in uncertain industries and technologies, the return to an investment consists of both the immediate return and an option value of learning. The option value is indirect value that arises when the knowledge gained from the investment help improve future investment decisions. An investment in a new unknown technology may have low expected immediate return, but a large option value, since success would spur additional investments in this technology. An investment in a well-known technology has low option value, since it is unlikely to alter investors' beliefs about the returns from the technology, and it has little effect on future investment decisions. Here, uncertainty is distinct from traditional risk and volatility, and it represents uncertainty about the underlying distribution of returns. A new and uncertain technology may turn out to be less risky than a known technology, but investors are initially uncertain about this, and the potential to learn generates the option value.

Option value creates an externality, and a social planner facing the same uncertainty will internalize this value. However, in a market where information is shared among investors, a free-rider problem reduces investors' incentives to explore and learn. With informational spillovers, an investor prefers the investment with the greatest immediate return, and will leave other investors to experiment with new, uncertain investments, knowing he can benefit from these, should they turn out successful (see

Bolton and Harris (1999) and Keller, Rady and Cripps (2005) for theoretical analyses of this problem).[2]

Uncertainty and option value are particularly large when investing in entrepreneurial companies. VCs are prominent investors in these companies and share a number of organizational features (see Sahlman (1990)) that allow them to internalize the option value: VCs are long-term investors and are compensated for long-term performance; they repeatedly invest in long-term projects with great uncertainty; they are in close contact with entrepreneurs, facilitating learning; and they are private investors in privately held companies, limiting informational spillovers.[3] In total, these features suggest that there is substantial option value when investing in entrepreneurial companies, and that VCs may be able to avoid the free-rider problem and internalize this value. The close relationship between the VC and the entrepreneur promote learning about new technologies in an uncertain environment; the long-term structure of the VC fund and the relationship with limited partners promote forward-looking investment behavior and lead VCs to internalize gains from future investments; and the private nature of the funds reduce informational spillovers, and potentially reduce the free-rider problem.

The analysis first presents a simple formal model of learning. This model is based on the statistical Multi-Armed Bandit model, and the returns from investing consist both of immediate return and an option value of learning about future returns. The model is calibrated using a dataset with U.S. venture capital investments, and it is found that option values vary between 3% and 53% of the total value of investing. VCs' investment decisions appear to be guided by option value to a significant extent. When VCs are

---

[2] In addition to the free-rider effect, Bolton and Harris (1999) also find a counteracting "encouragement" effect. It arises when current experimentation encourages other agents to experiment in the future, and this future experimentation is valuable for the investor. Keller, Rady and Cripps (2005) present conditions where this effect disappears and only the free-rider effect remains.

[3] Recently, the Freedom of Information Act (FOIA) forced disclosure of detailed information about investments by VC firms with certain limited partners. This created a vigorous response from the VC community and may limit investment by these investors in future VC funds.

classified according to the weight they place on option value, as well as the weight they place on general investment trends or public market signals, and how “random” their investments are, the results indicate that VCs who place more weight on option value have better performance and that VCs who make more random investment decisions have lower returns.

The Multi-Armed Bandit model presents a stochastic control problem, dating back to Robbins (1952). The model captures a simple situation with forward looking learning. The name refers to a situation where a gambler faces a number of slot machines (one-armed bandits). The gambler is uncertain about the distributions of the payoffs from the machines, but learns about the distributions from repeated play, and his problem is to determine the optimal gambling strategy. The fundamental trade-off is between *exploiting* machines with known payoff distributions and *exploring* machines with uncertain payoffs, hoping they will turn out profitable. In the words of Berry and Fristedt (1985) (p. 5), “it may be wise to sacrifice some potential early payoff for the prospect of gaining information that will allow for more informed choices later.” Whittle (1982) (p. 210) states that “[the Bandit problem] embodies in essential form a conflict evident in all human action. The “information versus immediate payoff” question makes the general problem difficult.”

In economics, the Bandit model has been used to model firms’ experimentation with prices to learn about uncertain demand (Rothschild (1974)). It has been widely applied in labor economics to model employee learning and job turnover (see i.e. Jovanovic (1979) and Miller (1984)). Weitzman (1979) considers a model of research projects and derives the optimal sequencing of these. In venture capital, Bergemann and Hege (1998) and Bergemann and Hege (2005) present theoretical models of staged financing based on the Bandit model. They address the question of how and for how long a VC should finance an entrepreneur given the potential for learning more about the quality of the entrepreneur’s project. Manso (2006) considers the closely related problem of giving an agent incentives to explore, and Bergemann and Valimaki (2006) briefly survey the Bandit literature.[4]

Methodologically, venture capital presents an attractive venue for studying learning and experimentation by firms. Unlike many other settings, data about VC investments contain individual projects (individual companies receiving financing) and the outcome of each of these. Projects are clearly delineated, and both their immediate returns and the learning process are apparent in the data. This contrasts other applications of the Bandit model, such as the application to job turnover. Typically, only job durations and changes are observed, and Heckman and Borjas (1980) argue that it is difficult to separate learning from individual heterogeneity in this case. Here, the agents' full learning history is observed and this identification problem is not a concern.

To keep the analysis tractable, the model is a simple model of forward-looking learning. As such, it presents a simplified view of VC investments. The model assumes a stationary environment[5] where investors learn from only their own investment histories. The model assumes that projects (entrepreneurial companies) arrive exogenously. Numerical tractability also imposes limitations on the estimation procedure. Option values are complex functions of investors' discount rates and prior beliefs and are difficult to calculate. As a result, the discount rate and prior beliefs are calibrated rather than estimated. More generally, empirical analysis of entrepreneurial finance always faces data limitations. Entrepreneurial companies, by definition, have short operating and financial histories and little information is systematically observed about these companies. This means that classifications of companies and investment outcomes are necessarily crude, although common in the literature. The paper should be read with these

---

[4] There is a substantial empirical literature about learning in general (see i.e. Crawford and Shum (2005), Erdem and Keane (1996) and Hitsch (2006) for learning about consumer demand). In this literature, the option value is subsumed into a general value function of a dynamic programming problem, and it is not separately quantified.

[5] Axelson, Stromberg and Weisbach (2006) contains a recent theoretical investigation of the cyclicality of the VC industry, and Gompers, Kovner, Lerner and Scharfstein (2005) investigate responses by different VCs to public market signals about shifts in fundamentals.

limitations in mind. The results are encouraging, and hopefully future extensions can address remaining shortcomings.

## I. The Multi-Armed Bandit Model

The model presented here is a Bandit model with infinite horizon, geometric discounting, and independent Bernoulli arms.[6] In this model, a single VC invests in one project each period. The opportunity cost of investing in a project in an unknown industry is that it prevents an investment in an industry with a certain return. The environment is stationary, and only investors' beliefs about returns from different industries evolve over time. The outcome of an investment is either success or failure, it is immediately observed, and the VC's beliefs are updated accordingly, using Bayes rule.

### *A. The Formal Model*

The investor faces an infinite sequence of periods, $t = 0,1, \dots$ . At each date $t$, the investor chooses between $K$ arms (investments in different industries). This choice is denoted $i(t) \in \{1,...,K\}$ . An investment in industry $i$ at time $t$ is either successful or not, as given by the random variable $y_i(t) \in \{0,1\}$ . The success probability is $p_i = \Pr[y_i(t) = 1]$ . The investor is uncertain about this probability, but has prior beliefs about $p_i$ , denoted $F_i(0)$ . The beliefs before investing at time $t$ are denoted $F_i(t)$ , and they are a function of the initial beliefs and the history of investment decisions and outcomes up to time $t$.

The investor's strategy specifies the investment decision as a function of his history, where the history consists of past investment decisions and outcomes. Hence, the strategy at time $t$ is a function $s(t) : \{1,...,K\}^t \times \{0,1\}^t \to \{1,...,K\}$ , and the full investment strategy is $S = \left\{ s(0), s(1), s(2), ... \right\}$ .

[6] See Berry and Fristedt (1985), Gittins (1989), and Whittle (1982) for general discussions of the Bandit model and related models and results.

The investor's problem is to choose the investment strategy that maximizes expected return. Let $\delta$ be the discount rate, and the investor's problem is

$$V = \sup_{S} E\left[ \sum_{t=0}^{\infty} \delta^t y_{s(t)}(t) \middle| F(0) \right] \tag{1}$$

The Bellman equation for this dynamic programming problem is

$$V\left(F(t)\right) = \max_{i=1,2,\ldots,K} E[y_i(t) \mid F_i(t)] + \delta E[V(F(t+1)) \mid F(t), i] \tag{2}$$

The investor's beliefs are state variables and they develop according to the transition rule

$$F_j(t+1) = F_j(t) \ for \ i(t) \neq j \tag{3}$$

$$F_j(t+1)(s) = \begin{cases} \dfrac{\int_0^s u f_j(t)(u)du}{\int_0^1 u f_j(t)(u)du} & for \ i(t) = j \ and \ \ y_j(t) = 1 \\[2ex] \dfrac{\int_0^s (1-u) f_j(t)(u)du}{\int_0^1 (1-u) f_j(t)(u)du} & for \ i(t) = j \ \ and \ \ y_j(t) = 0 \end{cases} \tag{4}$$

Equation (3) states that beliefs are only updated for the industry in which the investor invests, and remain unchanged for the remaining industries. Equation (4) reflects Bayesian updating of the investor's beliefs about the distribution of $p_i$ after choosing $i(t) = j$ and observing either $y_j(t) = 1$ or $y_j(t) = 0$.

The Bellman formulation illustrates the trade-off in this model. The first term in the maximization in Equation (2) gives the investor's expected immediate return from investing in industry $i$, and if there were no learning, this would be the entire return from the investment. The second term is the continuation value. Without learning the continuation value would be independent of the investment decision, but if the current investment decision may affect future investments, the continuation value is also a function of the investment decision. It increases with more informative investments, and this generates the additional option value of investing.

*B. The Gittins Index*

Gittins and Jones (1974) solve the general version of the investor's problem, and formulate the solution in terms of the *Dynamic Allocation index*, now called the *Gittins index*. For each industry, the Gittins index is calculated from the history of investments in just this industry, and the optimal strategy is to invest in the industry with the highest value of the index. In other words, if $v_i(t)$ is the Gittins index for industry *i*, at time *t*, the optimal strategy is

$$s(t) = \arg\max_{i=1,\dots,K} v_i(t) \tag{5}$$

To calculate the index, let $\tau$ denote a stopping time for industry *i*, and Gittins (1979) shows that the value of the index can be written as

$$v_i(t) = \sup_{\tau} \left\{ \frac{E\left[\sum_{s=t}^{\tau} \delta^s y_i(t) \mid F_i(t)\right]}{E\left[\sum_{s=t}^{\tau} \delta^s \mid F_i(t)\right]} \right\} \tag{6}$$

Calculating the index is numerically difficult,[7] but Gittins and Jones (1979) derive an illustrative approximation for the case where $\delta = .75$ and prior beliefs follow $Beta(a_{i,0}, b_{i,0})$ distributions.[8] The assumption of a Beta distributed prior is not required in the general Bandit model, but it is important for keeping the analysis tractable, and it is maintained below. With Beta distributed priors, Bayes theorem implies that the updated

[7] A MatLab program that calculates the Gittins index is available from the author.

[8] The *Beta*(*a*,*b*) distribution has density $f(s) = \frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)} s^{a-1}(1-s)^{b-1}$ for $a > 0$ and $b > 0$, and its mean is $a/(a+b)$.

beliefs are distributed $Beta(a_i, b_i)$, with $a_i \equiv a_{i,0} + r_i$ and $b_i \equiv b_{i,0} + n_i - r_i$. Here $r_i$ is the number of past successes, and $n_i$ is the total number of investments in the industry. As $a_i$ and $b_i$ increase, the mass of this distribution becomes concentrated at $\lambda_i \equiv a_i / (a_i + b_i)$, which equals the mean of the *Beta*($a_i$, $b_i$) distribution. Gittins and Jones (1979) then derive the approximation

$$v(a_i, b_i) \approx \lambda_i + \frac{1}{A(\lambda_i) + B(\lambda_i)(a_i + b_i)} \tag{7}$$

where $A(\lambda_i)$ and $B(\lambda_i)$ are two non-negative tabulated functions.

This approximation has an intuitive economic interpretation. The total value of investing in industry *i* is $v(a_i, b_i)$. The expected immediate return is $\lambda_i$, since, with a binary outcome and the value of success normalized to one, expected success probability, $\lambda_i$, equals the expected return, i.e. $E[y_i \mid F_i(t)] = \Pr[y_i = 1 \mid F_i(t)] = E[p_i \mid F_i(t)] = \lambda_i$. Clearly, the total value of investing is at least this big. The second term in equation (7) is the value of the investment in excess of the immediate return, and this represents the option value of learning. In the denominator, $(a_i + b_i) = a_{0,i} + b_{0,i} + n_i$, and for any fixed $\lambda_i$, the denominator grows in $n_i$ and the option value tends to zero. Economically, as the number of past investments increases, the beliefs become concentrated around $\lambda_i$, which approaches the true $p_i$, and the option value of learning vanishes.

### *C. Discussion of Index Result*

The assumption of Beta distributed priors and the index result simplify the state-space of the problem and make it empirically tractable. In principle, the Bandit problem is just another dynamic programming problem, where the posterior beliefs are the state variables. In theory, it could be solved by standard methods, such as iterating the value or policy functions.[9] However, these methods are only numerically tractable for problems with fairly small state spaces.[10] The assumption that the prior beliefs follow Beta distributions reduces the state space to two dimensions per arm. With arbitrary priors, the distributions of posterior beliefs would also be arbitrary, and the state space would be correspondingly high-dimensional (in principle, infinitely-dimensional). The Beta distribution is closed under Bayesian updating, and the posterior beliefs are fully characterized by its two parameters. With six industries, this leaves a twelve-dimensional state space, which is still numerically difficult.[11] The benefit of the index result is that the value of each arm can be calculated independently by solving the stopping problem from Equation (6). This divides the problem into six independent dynamic programming problems, each of these has a two-dimensional state space, which is quite tractable. For each investment, the option value can be calculated and included as a separate variable in the empirical analysis. This makes the analysis transparent and simple, and it is straightforward to scale the problem to an arbitrary number of arms.

## II. Data Description

### *A. Sample*

The data are provided by Sand Hill Econometrics. They contain the majority of VC investments in the U.S. in the period from 1987 to 2005.[12] The data extend two

---

[9] See Judd (1998) for a discussion of different methods for solving dynamic programming problems.

[10] Using these standard methods, Hitsch (2006) solves and estimates a related model involving a manufacturer experimenting and learning about the demand for ready-to-eat cereal products. This model has a four-dimensional state space.

[11] Erdem and Keane (1996) and Crawford and Shum (2005) estimate related learning models with state spaces comparable to the state space of the present model. Both papers rely on a certain interpolation technique to approximate the value function, and Erdem and Keane (1996) describe their estimation procedure as an "extremely computationally burdensome process."

[12] Sand Hill Econometrics has combined two existing commercial databases, Venture Xpert (formerly Venture Economics) and Venture Source (formerly Venture One) and has identified missing investments and verified and corrected the data, particularly information about companies' exit events.

commercial databases: Venture Xpert[13] and VentureOne. These databases have been extensively used in the VC literature (i.e. Kaplan and Schoar (2005) and Lerner (1995)). Gompers and Lerner (1999) and Kaplan, Sensoy and Strömberg (2002) investigate the completeness of the Venture Xpert data and find that they contain most VC investments, and that missing investments tend to be less significant ones.

The sample is constructed as follows. It is restricted to investments made up to 2000, since it typically takes VC backed companies four to five years to go public or to realize a return after the initial investment. It's common for multiple VCs to invest in a company, and the sample contains these different investments. However, when a VC firm invests in a company over multiple rounds, only the initial investment is included. It is necessary to remove VC firms making less than 40 investments, since their short investment histories make it difficult to draw inference about their learning, and this creates convergence problems for the estimation procedure. This removes 50% of the companies and 49% of the investments from the sample. The remaining investors represent the more active VC investors, which, not surprisingly, have higher success rates than the removed investors.[14] The use of this restricted sample does not introduce any immediate biases, but the estimates should be interpreted as estimates for the "top tier" investors. The final sample contains 19,166 investments in 6,076 companies by 216 VC firms.

### *B. Variables*

The main variables are the sequence of the VCs' investments in companies, the outcomes of these, and the industry classifications of the companies. These variables are used to construct the VCs' investment and learning histories. Summary statistics are in Table I.

---

[13] Venture Xpert is formerly known as Venture Economics.
[14] The eliminated VCs have *Success Rate*, *IPO Rate*, and *ACQ Rate* of 39%, 13%, and 26% respectively. For the remaining investors, the corresponding rates are 50%, 20%, and 30%.

Each company is classified in one of six industry classifications. These are: "Health / Biotech," "Communications / Media," "Computer Hardware / Electronics," "Consumer / Retail," "Software," and "Other." The corresponding control variables are: *Health*, *Communications*, *Computers*, *Consumer*, *Software*, and *Other*. These six broad industry classifications are aggregated from 25 minor classifications. The aggregation is necessarily somewhat arbitrary, but the intention is to classify companies into industries where experience with one company in informative about investments in other companies in the same industry.

For each investment, the outcome is given by the binary variable *Success*, and for each investor, *Success Rate* measures the investor's performance as the number of successful investments divided by total investments for this investor. An investment is successful, if the company subsequently goes public or is acquired. VC investors realize most of their return from a few successful investments, yet this measure is obviously a coarse measure of investment outcomes. Ideally, success would be measured in dollars or as a percentage return. However, data limitations prevent this, and the coarser measure is common in the literature. One potential concern is that investments in companies that are later acquired as part of their liquidation are counted as successful investments. Gompers and Lerner (2000) compare the broader measure to a number of other outcome measures (including just counting IPOs as successful) and find qualitatively small differences. The results here are also robust to using just using IPO as the success criterion.

Companies are classified as being early-stage or late-stage, where late-stage roughly corresponds to companies having regular revenues. The binary variable *Stage* equals one for late-stage companies, and 28.7% of investments in the sample are in late-stage companies.

Two variables capture public market signals and general trends in the market, such as the general shift towards communication and computer related investments in the late nineties. *Industry Investments* contains the total number of VC investments in each industry in each year in the data. It varies from a low of 36 investments in "Other" in 1994 to a high of 3,443 investments in "Computers" in 2000. Following Gompers, Kovner, Lerner and Scharfstein (2005), the variable *Industry IPOs* contains the number of IPOs each year for each industry. They find this variable is an important determinant of VCs' investments, particularly for VCs with more industry specific experience.

*C. Option Values and Expected Returns*

For each investor, at the time of each investment, the expected return and option value are calculated for each of the six industries. Note that expected return, $\lambda$, and *Option Value* depend on the investor's beliefs about $p_i$, but not on its true value, and it is *not* assumed that $p_i$ is constant across investors or industries. The calculation of the option values requires assumptions about investors' prior beliefs about $p_i$. These are taken to follow $Beta(a_{i,0}, b_{i,0})$ distributions with $a_{i,0} = 1$ and $b_{i,0} = 19$. This can be interpreted as the investor having experienced one previous success and 19 previous failures in each industry, and the investor initially expects a return of $\lambda_{i,j,t} = 1/20 = 5\%$. It is necessary to specify a low prior $\lambda$ relative to the observed success rate, to explain the degree of VC specialization and persistence in their investment decisions.[15] If $\lambda$ were equal to the empirical success rate, option values of untried industries would always lead investors to shift to these after investing in one industry and realizing the empirical success rate. The assumption implies that an investor who, say, is entirely specialized in biotech investments has fairly pessimistic beliefs about his potential returns from other industries, which seems reasonable.[16] Let $r_{i,j}$ be the number of past successes and $n_{i,j}$ be the total number of investments for investor $j$ in industry $i$, and these two variables summarizes the investor's investment history. Let $a_{i,j} \equiv a_{i,0} + r_{i,j}$ and $b_{i,j} \equiv b_{i,0} + n_{i,j} - r_{i,j}$, and from

[15] This also presents a simple counterexample to the claim by Erdem and Keane (1996), repeated by Crawford and Shum (2005), that "it is difficult to generate substantial persistence in choices without risk-aversion." In fact, in the limit, the Bandit model (with risk-neutral agents) always exhibit persistence (see Rothschild (1974)). Gompers, Kovner, Lerner and Scharfstein (2006) study persistence and specialization in venture capital investments. .

[16] In addition, the model is estimated using each investor's ten initial investments for "burn-in" of their beliefs without including them in the estimation. This leaves the results largely unchanged, although their statistical significance decreases somewhat due to the discarded observations. The model is also estimated with industry fixed effects to capture systematic differences in investors' beliefs across industries.

Bayes rule it follows that investor's updated beliefs about $p_i$ are distributed $Beta(a_{i,j}, b_{i,j})$. The investor's expected success rate is $\lambda_{i,j,t} = a_{i,j,t} / (a_{i,j,t} + b_{i,j,t})$. In the data, this variable varies between .03 and .72.

The discount factor is set to $\delta = .99$.[17] The average time between investments is 48 days and $\delta = .99$ then corresponds to an annual discount rate of 8%. The results are qualitatively similar for discount factors of $\delta = .75$[18] or $\delta = .95$, which correspond to annual rates of 783% and 47%, respectively. The option values are then calculated using an algorithm in Gittins (1989). In the data, *Option Value* varies from .06 to .08, and as a fraction of total value ($Option\ Value / (Option\ Value + \lambda)$), it varies from .03 to .53 with an average of .25. In other words, on average the option value is 25% of the value of an investment, but it can be small as 3% or as large as 53%.

**** TABLE I ABOUT HERE ****

## III. Empirical Results

At time *t*, the value for investor *j*, from investing in industry *i*, is specified as

$$v_{i,j,t} = \lambda_{i,j,t}\beta_1 + Option\ Value_{i,j,t}\beta_2 + X'_{i,j,t}\beta_3 + \varepsilon_{i,j,t} \tag{8}$$

where $\lambda_{i,j,t}$ is immediate expected return, $Option\ Value_{i,j,t}$ is option value, and $X_{i,j,t}$ contain additional control variables. The construction of the variables is described above. The investor chooses the investment with the highest total value, and when $\varepsilon_{i,j,t}$ follows an Extreme Value distribution, the probability of observing investor *j* investing in industry $i'$ is

[17] In the dynamic learning literature, it is common to set the discount rate, due to the difficulty of estimating it. Erdem and Keane (1996) set $\delta = .995$ (for weekly data, implying $r$ = 30%), Crawford and Shum (2005) set $\delta = .95$ (implying $r$ = 5%).

[18] An earlier draft of the paper uses this discount rate and the approximation of the Gittins index from Gittins and Jones (1979). The results are largely similar.

$$\Pr[i_{j,t} = i' \mid F_{i,j,t}] = \frac{\exp(v_{i',j,t})}{\sum_{i=1,\dots K} \exp(v_{i,j,t})} \tag{9}$$

These probabilities form the basis for the likelihood function, and the model is equivalent to the Multinomial Logit model (see McFadden (1973) and McFadden (1974)).

### *A. Evidence of Learning*

The specification in Equation (8) describes the investment decision as a function of immediate return and option value. Estimated coefficients are reported in Table II. In Specification 1, the estimated coefficients on both $\lambda_{i,j,t}$ and $\textit{Option Value}_{i,j,t}$ are positive and significant. Not surprisingly, investors prefer industries where they expect a higher immediate return. But, investors also prefer industries with higher option value. Their investment decisions appear to be forward looking, and the investors appear to appreciate the option value of learning. Investors seem to internalize the value of learning and actively engage in exploring new investments to learn about their returns.

**** TABLE II ABOUT HERE ****

The other specifications control for alternative factors that may affect investment decisions. Although not part of the formal model, investors also learn from other investors and general public market signals (see Gompers, Kovner, Lerner and Scharfstein (2005)). Specifications 2 and 3 control for the total number of VC backed IPOs and the number of VC investments in the six industries in each year. These two variables have small but positive and significant effects on investment decisions, consistent with investors being affected by overall market movements. But, including these effects does not eliminate the effects of $\lambda_{i,j,t}$ and $\textit{Option Value}_{i,j,t}$, and even after controlling for the general trends in the market, the investors still appreciate the value of learning from their own investments and explore as described by the Bandit model.

Investors may also be inherently timid, or restricted to investing in a few industries with a cost of switching to other industries.[19] In this case, option value would (inversely) proxy for the number of past investments the VC has made in an industry. To test this hypothesis, the variables *Industry Experience* and *Previous* are included in the estimation. *Industry Experience* contains the number of investments each investor has made in each industry at the time of each investment, and *Previous* is a dummy variable that equals one for the industry of the investor's previous investment. The estimated coefficients are positive and significant, but the coefficients on $\lambda_{i,j,t}$ and $Option\ Value_{i,j,t}$ remain positive and significant. Overall, this is consistent with the learning model. When investors receive positive information about an industry, they keep investing there, and *Industry Experience* and *Previous* will have positive significant coefficients. In the Bandit model, this updating is captured by $\lambda_{i,j,t}$ and $Option\ Value_{i,j,t}$, and the significance of these variables diminishes when *Industry Experience* and *Previous* are included. But again, $\lambda_{i,j,t}$ and $Option\ Value_{i,j,t}$ remain significant, and the learning model is not rejected. The final specification is a "kitchen-sink" regression with all regressors. The results remain largely unchanged.[20]

In the first specification in Table II, the coefficient on $\lambda_{i,j,t}$ is 4.7657 and the estimated coefficient on $Option\ Value_{i,j,t}$ is 6.4545. A literal reading of the Bandit model says that these two coefficients should be equal, since the immediate payoff and the option value are equally valuable to investors. The greater coefficient on $Option\ Value_{i,j,t}$ suggests that the model underestimates the option value or, equivalently, that investors explore more than predicted by the model. The coefficient for *Option Value* is greater than the coefficient on the immediate return across most specifications, and there are a

---

[19] Banks and Sundaram (1994) study a Bandit model with switching costs, and find that the index result does not extend to this case.
[20] Note that it is not possible to include time fixed effects in this model (alternatively, one can argue that the model is robust to time specific effects). In the multinomial model, a time specific effect can only enter through its interaction with other regressors.

number of possible explanations for this. First, it may be that actual prior beliefs are more dispersed than the priors specified here. This would lead the model to underestimate option value. In Table II, Specification 5 eliminates some of these differences by introducing industry specific effects, but the difference remains. Second, if investors' discount rates are smaller ($\delta$ closer to one) than assumed here, investors will value future returns more, and the model will underestimate option value. Generally, the results are robust to different choices of $\delta$, although there is a clear inverse relationship between $\delta$ and the estimated coefficient on *Option Value*. Third, the model may be misspecified. Investors may have access to learning or knowledge outside the model, the environment may not be stationary, or VCs may simply be acting suboptimally. It would be interesting distinguish these explanations, but this would add substantial theoretical and numerical complexity, and goes beyond the present scope. The argument here does not hinge on this distinction.

*B. Classifying Investment Strategies*

The model is now estimated for each investor separately. The value investing in industry *i* for investor *j* at time *t* is given by[21]

$$v_{i,j,t} = \left[\lambda_{i,j,t} + Opt\,Val_{i,j,t}\right]1 + Opt\;Val_{i,j,t}\,\gamma_{j,1} + Industry\,IPOs_{j,t}\,\gamma_{j,2} + \varepsilon_{i,j,t} \qquad (10)$$

The bracket contains the sum of the immediate return and the option value, and this is the Gittins index. The second term is option value in excess of the option value in the bracket, and $\gamma_{j,1}$ captures whether a VC assigns a greater or smaller weight to this value than predicted by the model. An investor with positive $\gamma_{j,1}$ has a more explorative investment behavior, and $\gamma_{j,1}$ classifies investors according to how explorative their

[21] In contrast to the specification in equation (8), this equation is estimated without industry specific effects. Since the coefficients are investor specific, adding five industry specific effects per investor would add 216 x 5 = 1080 additional coefficients. This substantially reduces the statistical power of the analysis.

investment strategies are. The coefficient $\gamma_{j,2}$ classifies investors according to how much they follow public market signals, here measured by *Industry IPOs*.

One unusual feature of this specification is that the scale of the equation is normalized by setting the "coefficient" for the term in the bracket to one. Usually, the scale is normalized by fixing the variance of the error term. The bracket is the total value of investing (the Gittins index), and normalizing the coefficient on this term makes the estimated coefficients comparable across investors. In addition, it allows for estimating the standard deviation of $\varepsilon_j$ for each investor. This coefficient, $\sigma_j$, measures how "random" or "opportunistic" the investor's strategy is.

To estimate the coefficients, a standard multinomial discrete choice model is estimated, and the coefficients are rescaled by the value of the first coefficient. This has the advantage that the standard error of the first coefficient provides a measure of how precisely the investor's coefficients are estimated. For investors with shorter investment, it is difficult to measure how they learn, and the estimates of the characteristics of their investment strategy, $\gamma_{j,1}$, $\gamma_{j,2}$, and $\sigma_j$, are imprecise. Below, investors are weighted according to the precision of these characteristics, and less weight is placed on investors with less precise coefficients. There is one disadvantage to this estimation procedure. A small number of investors have negative estimates of the first coefficient. These are typically investors with short investment histories and imprecisely estimated characteristics, but these investors show up with negative values of $\sigma_j$, which is unfortunate. Since these investors typically have low weights, all results are robust to including or excluding these investors, as well as replacing $\sigma_j$ by its absolute value.

In short, investment strategies are now classified along three dimensions. An investor with a high value of $\gamma_{j,1}$ places relatively more weight on forward looking learning and exploration. A high value of $\gamma_{j,2}$ is an investment strategy with more weight on overall market trends. A high value of $\sigma_j$ corresponds to an investment strategy that is more "random" or "opportunistic." There is one estimate of these three characteristics for each investor. To give the characteristics meaningful units, they are rescaled to make their standard deviation across investors equal to one (Table I, Panel B presents both scaled and raw estimates). In regressions, their coefficients are interpreted as the effect of a one standard deviation change in investment behavior within the sample of investors.

*C. Investment Strategy and Performance*

Consider now how each VC's performance depends on the investment strategy, as characterized above. Here, an investor's performance is measured by *Success Rate*, which is the number of successful investments over total investments. Using each of the 216 investors as a separate observation, the following regression is estimated

$$Success\ Rate_j = \beta_0 + (\gamma_{j,1})\beta_1 + (\gamma_{j,2})\beta_2 + (\sigma_j)\beta_3 + \varepsilon_j \qquad (11)$$

A positive estimate of $\beta_1$ indicates that investors that place more weight on *Option Value* (i.e. explore more) tend to perform better, a positive estimate of $\beta_2$ indicates that investors that place more weight on *Industry IPOs* (i.e. follow the general market more) perform better, and a positive estimate of $\beta_3$ indicates that investors with a higher standard error in the estimate of equation (10) (i.e. make more "random" or "opportunistic" investments) perform better. Estimates are in Table III, Panel A.

**** TABLE III ABOUT HERE ***

The first specification characterizes investment strategies in terms of $\gamma_{j,1}$ and $\sigma_j$. Investors with higher $\gamma_{j,1}$ have more explorative investments strategies, and have higher success rates. Investors with higher $\sigma_j$ make more random investments and have lower success rates. The magnitudes are economically meaningful. A one standard deviation increase in exploration is associated with a 2.14% to a 2.62% increase in success rate. A one standard deviation in "randomness," is associated with a drop in success rate between 1.60% and 2.40%.

The second specification includes $\gamma_{j,2}$, which captures investors' tendency to invest in industries with a greater number of VC backed IPOs. The coefficient is positive but insignificant, suggesting that investors that follow general trends more have slightly higher returns, although the evidence is weak. Specification 3 includes total number of investments by the investor (*Total Experience*). One would expect that investors with higher success rates make more investments and have larger experience. The estimated coefficient is positive, but insignificant, and there is little support for this relationship after controlling for the other characteristics of the investors' strategies.

Consider how the performance of the company is related to the investor's strategy. Panel B in Table III reports estimated coefficients from a Probit model where the probability of success for each investment is a function of the strategy and additional controls.

$$\Pr(IPO_{i,j,t}=1)=\Phi\left(\beta_0+(\gamma_{1,j})\beta_1+(\gamma_{2,j})\beta_2+(\sigma_j)\beta_3+X'_{i,j,t}\beta_4\right) \tag{12}$$

In Panel B, Specification 1 shows that investments made by more explorative investors are more likely to be successful, and investments made by more opportunistic investors are less likely to be successful. Again, the economic effects are meaningful. A one standard deviation increase in $\gamma_{1,j}$ increases the success probability by 1.48% to 3.35%. A one standard deviation increase in $\sigma_j$ decreases the success probability by 2.12% to 2.99%. Specification 2 includes investors' tendency to follow general market trends. Again the effect of $\gamma_{2,j}$ is small and insignificant. The final specification is the "kitchen-sink" regression with additional controls and fixed effects. Investments in companies at the late-stage are 15.14% more likely to be successful; investments by more experienced investors are marginally more likely to be successful; and investments in industries with many VC backed IPOs are marginally less successful. Overall, the results at the company level supports the evidence at the investor level, although the sign on $\gamma_{2,j}$ reverses.

*D. Speed of Investing*

It is interesting to consider how learning affects the speed of investments. It is possible that investors will first try an explorative investment in a new industry, and if it turns out successful, will accelerate the pace. Alternatively, investors may initially make a quick number of explorative investments, and then slowly settle on the sectors that appear most promising. The Bandit model does not explicitly incorporate speed, but it provides a simple framework for considering this question. In the model, speed is determined by the discount factor, $\delta$. The closer the factor is to one, the smaller the discounting, and the higher the speed. When investments have positive NPV, as here, investors would like to invest as fast as possible. So, as a starting point, assume that increasing speed requires costly effort. This cost may reflect a cost of more quickly searching for new investment opportunities, or a cost of accepting lower quality investments and working harder to improve them. Write the investor's problem as

$$V(F(t)) = \max_{i.e} E[y_i(t) \mid F(t)] - C(e) + \delta(e) E\left[V(F(t+1)) \mid F(t), i\right] \tag{13}$$

where $e$ is the investor's effort, $C(e)$ is an increasing convex cost of providing effort, and $\delta(e)$ is the discount rate, which tends to one as more effort increases the speed of investing and shortens the duration of a period. When the continuation value (i.e. the last term in Equation (13)) increases, the benefit of higher speed increases and the investor exerts more effort. Intuitively, when the value of future investments is high, regardless of whether this reflects option value or expected return, investments accelerate. This can be verified in the data, and the coefficients of the following regression are reported in Table IV.

$$Time\,to\,Investment_{i,j,t} = \beta_0 + Gittins_{i,j,t}\beta_1 + \varepsilon_{i,j,t} \tag{14}$$

Here *Time to Investment* is the number of days since the previous investment by the same investor. In the sample *Time to Investment* is on average 48.4 days with a standard deviation is 106.5. The continuation value is captured by the Gittins index.[22]

In the first specification in Table IV, the coefficient on Gittins is -93. Consistent with the above argument, investments with a higher Gittins index (i.e. a higher continuation value) are made quicker. In Specification 2, a number of industry and year controls are included, along with the investor's experience, and the effect persists. More experienced investors also appear to invest faster, although the economic magnitude is small.

In the remaining specifications, the option value and the immediate return enter separately. In the baseline model in Equation (13), they should have similar coefficients. However, *Option Value* has a positive effect and *Lamda* has a negative effect on the time to investment. Investors appear to wait longer before making more explorative investments with higher option value. When these investments turn out to be successful and the expected return increases, the negative coefficient on $\lambda$ indicates that the investment speed increases. One possible explanation is that learning not only affects the investor's beliefs, but also reduces the cost of investing, given by the $C(e)$ function. A skilled investor, with a particularly high $\lambda$ for an industry, would then find it optimal to invest more quickly. But, this result is sensitive to the specification. In Specification 4 and 5, more controls are included, and the coefficient on *Option Value* first becomes insignificant and then changes sign. In short, the baseline specification is consistent with the predictions of the learning model, but the distinction between the separate effects of exploration and exploitation on speed is less conclusive. A further refinement of the analysis would proceed by explicitly incorporating these effects into the investor's problem.

---

[22] Formally, the continuation value is the Gittins index scaled by a factor, see Whittle (1982), p. 214. Note also that the formal solution to the problem with effort must modify the continuation value to reflect future effort and its cost. This problem is not solved here, and the relationship is a conjecture, not a formal result.

*E. Additional Robustness Checks*

The main results are robust across a large number of unreported specifications. To examine the role of the discount rate, the model is estimated with $\delta = .75$, $\delta = .95$, and $\delta = .99$. The resulting estimates are qualitatively similar, although the option value is increasing in $\delta$ and the coefficient on option value correspondingly decreases as $\delta$ tends to one. In this sense, $\delta = .99$ is the most conservative of the three figures, resulting in the smallest coefficients on option value.

The model is estimated with two different definitions of a successful investment outcome. The case reported above is the case where a successful investment is defined by the company subsequently going public or being acquired, but the results are robust to using a more limited success criterion where only companies going public are classified as successful. This may alleviate the concern that the broader criterion would count acquisitions that are really liquidations as successful investments.

To determine the sensitivity of the results to the specification of prior beliefs, the model is also estimated using each investor's initial ten investments to "burn-in" the beliefs, and then excluding them from the analysis. The results are largely unchanged, although the statistical significance is reduced somewhat due to the excluded observations. In addition, industry fixed effects can control for systematic differences in beliefs across industries.

Finally, it may be of concern that only few companies went public after 2000. Consequently, the model is estimated restricting the sample to end in 2000, 1998, 1996, 1994, and 1992, and all results are robust across these restricted samples. The signs and economic magnitudes of the main coefficients are unchanged, and although their statistical significance is somewhat reduced due to the smaller sample sizes, they remain statistically significant.

## IV. Conclusion – Summary and Extensions

With uncertainty about the returns from different investments, learning about different returns is valuable. In a fully informed market, a free-rider problem reduces the investors' ability to internalize this value and limits the amount of exploration and learning. VCs' organizational structure serves to reduce information spillover. This may be a way to reduce the free-rider problem, and permit VCs to internalize the option value to engage in socially valuable experimentation and learning. To empirically investigate this claim, a Multi-armed Bandit model is estimated, and the empirical results suggest that VCs' investment decisions are substantially affected by the option value of learning.

The leaning model has additional implications for venture capital investments. First, it allows for categorizing VCs' investment behavior as being more or less exploratory, more or less driven by public market signals, and more or less random or opportunistic. Consistent with the learning model, investors that are more exploratory have higher success rates. Strategies that are more random are associated with lower success rates, and the evidence is inconclusive for strategies that follow public market signals to a smaller or greater extent. Second, the model provides a framework for studying the speed of investments. Consistent with the model, it is found that investments are made more quickly when they have a higher continuation value.

The Bandit model has been widely studied in the theoretical literature and in the empirical labor literature. The benefit of applying this model to the market for venture capital is that the agents' full learning histories are observed, and this overcomes an identification problem inherent in the traditional applications. In addition, it provides an explicit dynamic model of VCs' investment decisions, and it presents an interesting theoretical and empirical setting for studying the evolution of these investors. The results are encouraging and further refinements of the methods introduced here are likely to lead to a better understanding of learning processes in venture capital and in other markets.

One interesting refinement would be to study learning at the finer level of a technology or an idea. Naturally, the current results are consistent with learning taking place at a finer level, and the current results provide clear evidence that investors do learn. But, it would be natural and interesting to refine the level of learning to the individual technology or idea. However, this presents a difficult challenge. When investment categories are defined at a finer level, there are few repeated investments in each category, and it becomes necessary to consider informational spillovers between related categories. In this case, the index result no longer holds, and this creates a numerically difficult dynamic programming problem. Solving this problem provides an interesting avenue for future research.

Finally, Gompers, Kovner, Lerner and Scharfstein (2005) find that changes in economic fundamentals as signaled by public market signals are important drivers of venture capital investments. The analysis here suggests that there are two fundamentally different kinds of changes to fundamentals that can make a new investment attractive. Either the expected immediate return can increase (an upward shift in $F(t)$) or the option value of the investment can increase (an increased "spread" in $F(t)$). For example, an increase in immediate return may arise from a new improvement in the production of a known product, i.e. following an investment in a technology that reduces its marginal costs. This has immediate value, but little new is learned about the demand of the product and it is unlikely to spur additional investments. In contrast, a new technology leading to a new product with entirely unknown applications and demand has high option value. If successful, it would cause additional investments and expansion in the application and supply of the product. Classical financial markets are well suited for responding to the first type of changes to fundamentals. However, the free-rider problem is inherent in the second type, and the argument made here suggests that the organizational structure of VCs make them particularly suited to respond to this kind of changes to fundamentals. Obviously, it is important that investments respond appropriately to changes in fundamentals, and exploring responses by different financial institutions to these changes may provide new insights into the role and value of venture capitalists in the economy.

**Bibliography**


Axelson, Ulf, Per Stromberg, and Michael Weisbach, 2006, The Financial Structure of Private Equity Funds, *working paper*.

Banks, Jeffrey, and Rangarajan Sundaram, 1994, Switching Costs and the Gittins Index, *Econometrica*, 62, 687-694.

Bergemann, Dirk, and Ulrich Hege, 1998, Dynamic Venture Capital Financing, Learning and Moral Hazard, *Journal of Banking and Finance*, 22, 703-735.

Bergemann, Dirk, and Ulrich Hege, 2005, The Financing of Innovation: Learning and Stopping, *RAND Journal of Economics*, 36.

Bergemann, Dirk, and Juuso Valimaki, 2006, Bandit Problems, *Cowles Foundation Discussion Paper, no. 1551*.

Berry, Donald, and Bert Fristedt, 1985. *Bandit problems: Sequential Allocation of Experiments* (Chapman and Hall, New York).

Bolton, Patrick, and Christopher Harris, 1999, Strategic Experimentation, *Econometrica*, 67, 349-374.

Crawford, Gregory, and Matthew Shum, 2005, Uncertainty and Learning in Pharmaceutical Demand, *Econometrica*, 73, 1137-1174.

Erdem, Tülin, and Michael Keane, 1996, Decision-Making Under Uncertainty: Capturing Dynamic Brand Choice Processes in Turbulent Consumer Goods Markets, *Marketing Science*, 15, 1-20.

Gittins, J. C., 1979, Bandit Processes and Dynamic Allocation Indices (with Discussion), *Journal of Royal Statistical Society. Series B (Methodological)*, 41, 148-164.

Gittins, John. C., 1989. *Multi-armed Bandit Allocation Indices* (John Wiley & Sons, New York).

Gittins, John. C., and D. M. Jones, eds., 1974. *A Dynamic Allocation Index for the Sequential Allocation of Experiments* (North-Holland, Amsterdam).

Gittins, John. C., and D. M. Jones, 1979, A Dynamic Allocation Index for the Discounted Multiarmed Bandit Problem, *Biometrika*, 66, 561-565.

Gompers, Paul, Anna Kovner, Josh Lerner, and David Scharfstein, 2005, Venture Capital Investment Cycles: The Impact of Public Markets, *working paper*.

Gompers, Paul, Anna Kovner, Josh Lerner, and David Scharfstein, 2006, Specialization and Success: Evidence from Venture Capital, *working paper*.

Gompers, Paul, and Josh Lerner, 1999. *The Venture Capital Cycle* (MIT Press, Cambridge).

Gompers, Paul, and Josh Lerner, eds., 2000. *The Determinants of Corporate Venture Capital Successes: Organizational Structure, Incentives, and Complementarities* (University of Chicago Press, Chicago, IL).

Gorman, M, and William Sahlman, 1989, What Do Venture Capitalists Do? *Journal of Business Venturing*, 4, 231-248.

Heckman, James, and George Borjas, 1980, Does Unemployment cause Future Unemployment? Definitions, Questions and Answers form a Continuous Time Model of Heterogeneity and State Dependence, *Econometrica*, 47, 247-83.

Hellmann, Thomas, and Manju Puri, 2000, The Interaction Between Product Market and Financing Strategy: The Role of Venture Capital, *The Review of Financial Studies*, 13, 959-984.

Hellmann, Thomas, and Manju Puri, 2002, Venture Capital and the Professionalization of Start-Up Firms: Empirical Evidence, *Journal of Finance*, 57, 169-197.

Hitsch, Günter, 2006, An Empirical Model of Optimal Dynamic Product Launch and Exit Under Demand Uncertainty, *Marketing Science*, 25, 25-50.

Jovanovic, Boyan, 1979, Job Matching and the Theory of Turnover, *Journal of Political Economy*, 87, 972-90.

Judd, Kenneth, 1998. *Numerical Methods in Economics* (MIT Press, Cambridge).

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

Kaplan, Steven, Berk Sensoy, and Per Strömberg, 2002, How Well do Venture Capital Databases Reflect Actual Investments? *working paper*.

Kaplan, Steven, and Per Strömberg, 2004, Characteristics, Contracts, and Actions: Evidence from Venture Capitalist Analyses, *Journal of Finance*, 59, 2177-2210.

Keller, G., S. Rady, and M. Cripps, 2005, Strategic Experimentation with Exponential Bandits, *Econometrica*, 73, 39-68.

Lerner, Josh, 1995, Venture Capitalists and the Oversight of Private Firms, *Journal of Finance*, 50, 301-318.

Manso, Gustavo, 2006, Motivating Innovation, *working paper*.

McFadden, Daniel, ed., 1973. *Conditional Logit Analysis of Qualitative Choice Behavior* (Academic Press, New York).

McFadden, Daniel, 1974, The Measurement of Urban Travel Demand, *Journal of Public Economics*, 3, 303-328.

Megginson, William, and Kathleen Weiss, 1991, Venture Capitalist Certification in Initial Public Offerings, *Journal of Finance*, 46, 879-903.

Miller, Robert A., 1984, Job Matching and Occupational Choice, *Journal of Political Economy*, 92, 1986-1120.

Robbins, H., 1952, Some Aspects of the Sequential Design of Experiments, *Bulletin of American Mathematical Society*, 55, 527-535.

Rothschild, M., 1974, A Two-Armed Bandit Theory of Market Pricing, *Journal of Economic Theory*, 9, 185-202.

Sahlman, William, 1990, The Structure and Governance of Venture Capital Organizations, *Journal of Financial Economics*, 27, 473-521.

Weitzman, M., 1979, Optimal Search for the Best Alternative, *Econometrica*, 47, 641-654.

Whittle, Peter, 1982. *Optimization Over Time: Dynamic Programming and Stochastic Control* (John Wiley & Sons, New York).

## TABLE I: Summary Statistics

**PANEL A: Summary Statistics By Company**

| | Obs. | Mean | Std. Dev. | Min | Max |
|---|---|---|---|---|---|
| IPO | 6,076 | 0.170 | 0.376 | 0 | 1 |
| Acquisition | 6,076 | 0.329 | 0.470 | 0 | 1 |
| Success (IPO + Acq) | 6,076 | 0.499 | 0.500 | 0 | 1 |
| Year | 6,076 | 1995.2 | 4.422 | 1987 | 2000 |
| | | | | | |
| Industry Classifications | | | | | |
| Health | 6,076 | 0.181 | 0.385 | 0 | 1 |
| Commmunications | 6,076 | 0.208 | 0.406 | 0 | 1 |
| Computers | 6,076 | 0.240 | 0.427 | 0 | 1 |
| Software | 6,076 | 0.175 | 0.380 | 0 | 1 |
| Consumer | 6,076 | 0.122 | 0.327 | 0 | 1 |
| Other | 6,076 | 0.074 | 0.261 | 0 | 1 |

**PANEL B: Summary Statistics by Investor**

| | Obs. | Mean | Std. Dev. | Min | Max |
|---|---|---|---|---|---|
| IPO Rate | 216 | 0.204 | 0.094 | 0.000 | 0.523 |
| Acq Rate | 216 | 0.299 | 0.067 | 0.106 | 0.492 |
| Success Rate | 216 | 0.503 | 0.118 | 0.133 | 0.864 |
| Total Experience | 216 | 88.731 | 64.210 | 40 | 577 |
| | | | | | |
| Classifications of investment strategy: | | | | | |
| Option Value | 216 | 2.762 | 27.551 | -80.801 | 226.649 |
| Standard Error | 216 | 0.292 | 0.608 | -4.213 | 4.640 |
| Industry IPOs | 216 | 0.002 | 0.016 | -0.172 | 0.056 |
| Normalized scale: | | | | | |
| Option Value | 216 | 0.100 | 1.000 | -2.933 | 8.226 |
| Standard Error | 216 | 0.480 | 1.000 | -6.933 | 7.634 |
| Industry IPOs | 216 | 0.115 | 1.000 | -11.040 | 3.602 |

**PANEL C: Summary Statistics by Investment**

| | Obs. | Mean | Std. Dev. | Min | Max |
|---|---|---|---|---|---|
| Experience | 19,166 | 68.081 | 73.230 | 1 | 577 |
| Stage | 19,166 | 0.287 | 0.452 | 0 | 1 |
| Year | 19,166 | 1995.2 | 4.542 | 1987 | 2000 |
| Success | 19,166 | 0.571 | 0.495 | 0 | 1 |
| Lambda | 19,166 | 0.260 | 0.153 | 0.032 | 0.717 |
| Option Value | 19,166 | 0.062 | 0.012 | 0.019 | 0.082 |
| Gittins Index | 19,166 | 0.323 | 0.147 | 0.063 | 0.745 |
| OptionValue / Gittins Index | 19,166 | 0.252 | 0.138 | 0.031 | 0.531 |

## TABLE I: Summary Statistics (cont.)

Panel D presents the number of VC investments and VC backed IPOs (in parenthesis) for each industry in each year.

**PANEL D: INVESTMENTS (IPOs) PER INDUSTRY PER YEAR**

| Year | Health | Comm | Comp | Cons | Soft | Other | Total |
|---|---|---|---|---|---|---|---|
| 1987 | 806 | 420 | 1,125 | 164 | 359 | 310 | 3,184 |
| | (3) | (2) | (4) | () | () | (2) | (11) |
| 1988 | 592 | 237 | 770 | 92 | 327 | 208 | 2,226 |
| | (4) | (4) | (5) | (1) | (1) | (4) | (19) |
| 1989 | 395 | 148 | 371 | 57 | 189 | 139 | 1,299 |
| | (11) | (1) | (7) | (4) | (4) | (5) | (32) |
| 1990 | 283 | 101 | 256 | 56 | 196 | 89 | 981 |
| | (11) | (4) | (10) | (1) | (6) | (4) | (36) |
| 1991 | 258 | 100 | 164 | 57 | 206 | 54 | 839 |
| | (45) | (11) | (14) | (2) | (8) | (3) | (83) |
| 1992 | 372 | 152 | 142 | 44 | 257 | 58 | 1025 |
| | (59) | (15) | (18) | (11) | (11) | (6) | (120) |
| 1993 | 357 | 159 | 123 | 74 | 163 | 57 | 933 |
| | (35) | (14) | (36) | (9) | (18) | (15) | (127) |
| 1994 | 338 | 176 | 159 | 65 | 189 | 36 | 963 |
| | (33) | (13) | (27) | (6) | (16) | (6) | (101) |
| 1995 | 445 | 240 | 191 | 134 | 280 | 86 | 1376 |
| | (40) | (17) | (30) | (5) | (33) | (6) | (131) |
| 1996 | 472 | 429 | 291 | 176 | 458 | 93 | 1919 |
| | (72) | (35) | (28) | (16) | (43) | (13) | (207) |
| 1997 | 621 | 533 | 391 | 269 | 633 | 124 | 2,571 |
| | (39) | (18) | (21) | (9) | (17) | (10) | (114) |
| 1998 | 688 | 723 | 426 | 410 | 739 | 231 | 3,217 |
| | (9) | (23) | (16) | (9) | (11) | (2) | (70) |
| 1999 | 844 | 1938 | 1055 | 1714 | 1249 | 175 | 6,975 |
| | (14) | (94) | (27) | (53) | (68) | (3) | (259) |
| 2000 | 961 | 2824 | 3443 | 1388 | 1034 | 359 | 10,009 |
| | (60) | (44) | (29) | (32) | (48) | (8) | (221) |
| Total | 7,432 | 8,180 | 8,907 | 4,700 | 6,279 | 2,019 | 37,517 |
| | (435) | (295) | (272) | (158) | (284) | (87) | (1,531) |

## TABLE II: Aggregate Investment Decisions

The table reports estimates of a Multinomial Logit model (McFadden choice model) where investors' industry choice is the endogenous variable. The possible choices are Health, Communications, Computers, Consumer Goods, Software, and Other. *Lambda* and *OptionValue* are investors' expected immediate return and option value of investing. *Industry Investments* is total number of investments in each industry per year across all investors in the data. *Industry Experience* is the past number of investments by the investor in the industry. *Previous* is a binary variable that equals one for the industry of the investor's previous investment. Observations are weighted according to the precision of the estimate of OptionValue (see text for details). Robust standard errors with clustering at the company level are in parenthesis. ***, **, and * denote statistical significance at the 1%, 5%, and 10% levels, respectively.

| | 1 | | 2 | | 3 | | 4 | | 5 | |
|---|---|---|---|---|---|---|---|---|---|---|
| Lambda | 4.7657 | *** | 4.2445 | *** | 4.3202 | *** | 3.3410 | *** | 2.4023 | *** |
| | (.0718) | | (.0788) | | (.0780) | | (.1354) | | (.1484) | |
| Option Value | 6.4545 | *** | 5.3043 | *** | 3.0443 | *** | 17.2952 | *** | 8.5117 | *** |
| | (.8884) | | (.9006) | | (.9344) | | (1.2284) | | (1.3057) | |
| Industry IPOs | | | 0.0085 | *** | | | | | 0.0188 | *** |
| | | | (.0004) | | | | | | (.0018) | |
| Industry Investments | | | | | 0.0007 | *** | | | 0.0006 | *** |
| | | | | | (.0000) | | | | (.0000) | |
| Industry Experience | | | | | | | 0.0228 | *** | 0.0166 | *** |
| | | | | | | | (.0018) | | (.0019) | |
| Previous | | | | | | | | | 0.2978 | *** |
| | | | | | | | | | (.0174) | |
| Industry Controls | No | | No | | No | | No | | Yes | |
| Observations | 19,166 | | 19,166 | | 19,166 | | 19,166 | | 19,166 | |

## TABLE III: Investment Strategies and Outcomes

Panel A shows estimated coefficients for an OLS regression. An observation is an investor and the endogenous variable is the investor's success rate. Panel B presents marginal effects estimated from a Probit model. Each observation is an investment in a company and the endogenous variable is the outcome. *Option Value*, *Standard Error*, and *Industry IPOs* characterize the investor's investment strategy in terms of its dependence on option value, its standard error, and on the number of VC backed IPOs in the industry in the same year. These coefficients are normalized to have standard error equal one (see text for details). *Experience* measures number of previous investments by the investor at the time of each investment. *Total Experience* is investor's experience at the end of the sample. *Stage* is an indicator variable that equals one for investments in late-stage companies. Observations are weighted according to the precision of the estimates (see text for details). Robust Standard errors with clustering at the company level are in parenthesis. ***, **, and * denote statistical significance at the 1%, 5%, and 10% levels, respectively.

**PANEL A: Success Rate of Venture Capital Firms**

| | 1 | | | 2 | | | 3 | | |
|---|---|---|---|---|---|---|---|---|---|
| | Coef. | Std. Err. | | Coef. | Std. Err. | | Coef. | Std. Err. | |
| Classification of Strategy | | | | | | | | | |
| Option Value | 0.0214 | (.0027) | *** | 0.0234 | (.0087) | *** | 0.0262 | (.0089) | *** |
| Standard Error | -0.0160 | (.0022) | *** | -0.0212 | (.0071) | *** | -0.0240 | (.0080) | *** |
| Industry IPOs | | | | 0.0085 | (.0112) | | 0.0086 | (.0117) | |
| Total Experience | | | | | | | 0.0001 | (.0001) | |
| Constant | 0.5432 | (.0044) | *** | 0.5441 | (.0094) | *** | 0.5285 | (.0126) | *** |
| Observations | 216 | | | 216 | | | 216 | | |

**PANEL B: Success of Individual Investments**

| | 1 | | | 2 | | | 3 | | |
|---|---|---|---|---|---|---|---|---|---|
| | dF/dX | Std. Err. | | dF/dX | Std. Err. | | dF/dX | Std. Err. | |
| Classification of Strategy | | | | | | | | | |
| Option Value | 0.0335 | (.0058) | *** | 0.0148 | (.0060) | *** | 0.0154 | (.0060) | *** |
| Standard Error | -0.0299 | (.0062) | *** | -0.0242 | (.0075) | *** | -0.0212 | (.0076) | *** |
| Industry IPOs | | | | 0.0029 | (.0094) | | -0.0013 | (.0095) | |
| Stage | | | | | | | 0.1514 | (.0157) | *** |
| Experience | | | | | | | 0.0001 | (.0001) | |
| Industry IPOs | | | | | | | -0.0002 | (.0005) | |
| Year Controls | No | | | Yes | | | Yes | | |
| Industry Controls | No | | | No | | | Yes | | |
| Observations | 19,166 | | | 19,166 | | | 19,166 | | |

## TABLE IV: Investment Speed

The table reports estimated coefficients from an OLS regression. Each observation is an investment by an investor, and the time to next investment (measured in days) is the endogenous variable. Gittins is the Gittins index of the investment, Option Value is the option value, and Lambda is the expected immediate return. Experience is the investor's experience, measured as the number of past investments. Standard errors clustered at the company level are reported in parenthesis. ***, **, and * denote statistical significance at the 1%, 5%, and 10% levels, respectively.

| | 1 | | 2 | | 3 | | 4 | | 5 | |
|---|---|---|---|---|---|---|---|---|---|---|
| | Coef. | Std. Err. | Coef. | Std. Err. | Coef. | Std. Err. | Coef. | Std. Err. | Coef. | Std. Err. |
| Gittins | -92.3023 | *** | -118.1745 | *** | | | | | | |
| | (5.7294) | | (8.8156) | | | | | | | |
| Option Value | | | | | 668.1704 | | 109.1238 | | -439.9838 | |
| | | | | | (66.0140) | *** | (86.5481) | | (92.8493) | *** |
| Lambda | | | | | -67.8572 | | -59.1082 | | -119.8716 | |
| | | | | | (6.7266) | *** | (6.6045) | *** | (9.0141) | *** |
| Experience | | | -0.0960 | *** | | | -0.1625 | | -0.1277 | |
| | | | (.0085) | | | | (.0118) | *** | (.0110) | *** |
| Constant | 78.3253 | *** | 84.7510 | *** | 24.5984 | | 82.7178 | | 105.6546 | |
| | (2.4732) | | (3.9776) | | (5.6587) | *** | (7.6471) | *** | (7.8387) | *** |
| Industry Controls | No | | Yes | | No | | Yes | | Yes | |
| Year Controls | No | | Yes | | No | | No | | Yes | |
| Observations | 18,950 | | 18,950 | | 18,950 | | 18,950 | | 18,950 | |