Learning by Investing: Evidence from Venture Capital

Morten Sorensen†

March, 2008

Abstract: Uncertainties about technologies and investment opportunities are prevalent for investments in entrepreneurial companies by venture capitalists (VCs), and this study finds that the resolution of these uncertainties, through VCs' learning, is important for their investment decisions. The hypothesis that individual investments are evaluated in isolation, as predicted by standard models, is clearly rejected. The empirical analysis is based on a dynamic learning model derived from the Multi-armed Bandit model. The results suggest that VCs learn from past investments (exploitation) but also consider the option value of future learning (exploration) when making investment decisions.

---

† Assistant Professor, University of Chicago GSB, and NBER, phone: 773-834-1726, email: msorense@ChicagoGSB.edu. I want to thank Dirk Bergemann, Jean-Pierre Dube, Thomas Hellmann, Gunter Hitsch, Boyan Jovanovic, Steve Kaplan, Josh Lerner, Laura Lindsey, Gustavo Manso, Lubos Pastor, Francisco Perez-Gonzalez, Josh Rauh, Bryan Routledge, and Amir Sufi for helpful suggestions and discussions. Arthur Korteweg provided excellent research assistance. I am grateful to participants at the 2007 AEA (Chicago), 2007 MFA (Minneapolis), and 2008 AFA (New Orleans) meetings, Arizona State University, CEBR (Copenhagen), London Business School, NBER Entrepreneurship workshop, University of Illinois-Chicago, and University of Illinois-Urbana Champaign for helpful comments. I want to thank Susan Woodward of Sand Hill Econometrics for generous access to data and the Kauffman Foundation for financial support.

Given their importance for financing high-tech entrepreneurial companies, venture capitalists (VCs) have substantial impact on innovation and development of new technologies (Kortum and Lerner (2000)). While a great deal has been written about the relationship between VCs and their portfolio companies,[1] less is known about VCs' decisions to invest in particular industries and companies.[2] These investments are fraught with uncertainty (Quindlen (2000)), and to understand VCs' investment decisions, it is important to understand how these uncertainties are resolved, i.e. how VCs learn. This study finds that VCs' investment decisions are affected both by the expected return from the investments themselves but also by the potential to learn from them. Learning is valuable, since it increases the investors' understanding of various investment opportunities and improves their future decisions. The hypothesis that VCs' investments are chosen independently to maximize the return from each investment individually, as predicted by standard models, is clearly rejected.

The empirical evidence is found using an empirical model derived from the Multi-armed Bandit model (see Berry and Fristedt (1985) and Gittins (1989)). The central ingredient in this model is the investors' beliefs about the profitability of their investment opportunities and the dynamics of these beliefs. Beliefs are shaped by their past investments and their outcomes, and they affect investment decisions in two ways. First,

---

[1] See, for example, Gorman and Sahlman (1989), Sahlman (1990), Lerner (1995), Gompers and Lerner (1999), Hellmann and Puri (2000), Hellmann and Puri (2002), Kaplan and Strömberg (2004), and Hochberg, Ljungqvist and Lu (2006).
[2] Notable exceptions are Kaplan and Strömberg (2004) and Gompers, Kovner, Lerner and Scharfstein (2005).

not surprisingly, investors prefer investments with greater expected immediate returns. The immediate return is the direct return from the investment itself. This is the return that is usually captured by economic models, and without learning, it is the only return. With learning, an additional indirect effect arises since investors prefer investments with more informative outcomes, because these investments help them learn and improve their future decisions. This value can be viewed as an option value of learning, and it generates a trade-off between *exploiting* investments with known payoffs and *exploring* investments with uncertain payoffs.[3] This study provides an empirical methodology for separating and measuring exploration and exploitation. To my knowledge, this is the first time these two effects have been estimated separately,[4] and the results confirm that both of them are important determinants of VCs' investment decisions.

Additional predictions from the model are that more valuable investments are made faster, and that VCs that explore more are more successful. These predictions are confirmed empirically, lending further support to the model. Methodologically, the study shows that the statistical *index result* can simplify the empirical analysis by formulating the sequential decision problem in terms of an econometric discrete choice model.

---

[3] The terminology of *exploitation* and *exploration* is introduced by March (1991) in the context of organizational learning.

[4] Previous applications of the bandit model include Rothschild (1974)'s model of firms' experimentation with prices to learn about uncertain demand and Weitzman (1979)'s analysis of optimal sequencing of research projects. Manso (2006) study incentive provision in a learning model similar to the Bandit model, and Bergemann and Hege (1998) and Bergemann and Hege (2005) present theories of staged financing based on the bandit model in the context of VC investments. Empirically, Jovanovic (1979) and Miller (1984) estimate models of job turnover in which workers learn about job-specific skills.

Previous studies of learning[5] use computationally intensive estimation procedures to capture the inherent dynamic programming problem (i.e. Crawford and Shum (2005) and Erdem and Keane (1996)). The index result leads to more transparent and tractable inference, allowing the model to be estimated using standard statistical procedures.

Taking the Bandit model to the data requires additional assumptions. As a starting point, the model assumes that investors chose between investments at the industry level, and that learning takes place at this level as well. This is a natural starting point. It is motivated both by data limitations and by other assumptions inherent in the Bandit model, and Goldfarb, Kirsch and Miller (2007) find evidence of VC behavior consistent with learning at the industry level. The data limitations arise since inference about learning behavior is derived from the VCs' investment histories. The model links their past investments and outcomes to their subsequent investment decisions, and identification is based on comparing investments in industries with shorter and longer histories, corresponding to greater and smaller potentials for learning. Defining categories at finer levels results in more individual categories and shorter histories in each of these, reducing the power of the empirical estimation procedure.

The model assumes that the environment is stationary, that investors only learn from their own past investments, and that investments in one industry are uninformative

---

[5] Another strand of the learning literature considers learning-by-doing, where learning is a free byproduct of an activity (Arrow (1962)). Recent finance applications include Pastor, Taylor and Veronesi (2006), Linnainmaa (2006), and Hochberg, Ljungqvist and Vissing-Jorgensen (2008), who study learning by the limited partners in VC funds.

about investments in other industries. These assumptions are technical assumptions required for the index result, and they are more reasonable at the industry level than at finer levels. Their severity can be mitigated somewhat by including market-level controls in the empirical analysis to control for general market trends and the effect of other investors' decisions. A more formal relaxation of these assumptions requires explicitly solving the investors' dynamic programming problems as part of the estimation procedure, making the inference problem largely intractable. With these assumptions, the index result overcomes this.

After investing in a company, the outcome of each investment is either a success (subsequent IPO or acquisition of the company) or a failure (liquidation). While this is a coarse outcome measure, it is difficult to obtain more detailed information about the financial performance of these investments, and this classification is standard in the literature. Binary outcomes also simplify they updating of investors' beliefs, since the resulting beliefs are easily captured by the Beta distribution.

The paper proceeds as follows. The following section presents the theoretical learning model and the index result. The second section presents the data and variables, and discusses the econometric implementation of the model. Section three presents the empirical evidence of learning. Section four discusses the construction and interpretation of the investors' prior beliefs, and the final section concludes.

I. The Multi-Armed Bandit Model

In the Multi-armed Bandit model[6] an investor faces an infinite sequence of periods, $t = 0, 1, \ldots$. Each period the investor chooses between K arms, denoted $i = 1, 2, \ldots, K$, where each arm represents an investment in an entrepreneurial company in industry i. The outcome of an investment is either a success or a failure, as indicated by $y_i(t) \in \{0,1\}$, and the success probability, given by $p_i = \Pr[y_i(t) = 1]$, is constant over time but can vary across industries and investors. The investor does not know p_i but has prior beliefs, given by $F_i(0)$. These beliefs are updated after each investment, using Bayes rule, and the updated beliefs before investing at time t are $F_i(t)$. The support of $F_i(t)$ is the interval from 0 to 1, representing all possible values of p_i.

With binary outcomes, it is convenient to assume that investors' initial beliefs are distributed $Beta(a_0, b_0)$. Let r_i be the number of past successes and n_i be the total number of past investments in industry I, and the updated beliefs are then $Beta(a_i, b_i)$, where $a_i = a_{i,0} + r_i$ and $b_i = b_{i,0} + n_i - r_i$. In other words, a_i counts the number of past successes and b_i counts the number of past failures. As the number of investments increases, the mass of the distribution becomes concentrated at the empirical success rate, defined as $\lambda_i = a_i / (a_i + b_i)$, which also equals the mean of the $Beta(a_i, b_i)$ distribution. This value

[6] See Berry and Fristedt (1985), Gittins (1989), and Whittle (1982) for general discussions of this class of models.

equals the investors' expected immediate return from the investment, since $E[y_i \mid F_i(t)] =$ $\Pr[y_i = 1 \mid F_i(t)] =$ $E[p_i \mid F_i(t)] = \lambda_i$.

The investor's strategy specifies investments as a function of past investments and their outcomes. The strategy at time *t* is $s(t): \{1,...,K\}^t \times \{0,1\}^t \to \{1,...,K\}$, and the full strategy is $S = \{s(0), s(1), s(2), ...\}$. The investor's problem is to determine the strategy that maximizes total expected return. Let δ denoted the discount factor, and the investor solves

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

Formulated as a dynamic programming problem, the Bellman equation is

$$V\left(F(t)\right) = \max_{s(t)=1,2,...,K} E\left[y_{s(t)}(t) \mid F(t)\right] + \delta E\left[V\left(F(t+1)\right) \mid F(t), s(t)\right]. \tag{2}$$

where the state variables contain the investor's updated beliefs. These develop according to the transition rules

$$F_i(t+1) = F_i(t) \; for \; s(t) \neq i \tag{3}$$

$$F_i(t+1)(v) = \begin{cases} Beta(a_i+1, b_i) & for \; s(t) = i \; and \; y_i(t) = 1 \\ Beta(a_i, b_i+1) & for \; s(t) = i \; and \; y_i(t) = 0 \end{cases} \tag{4}$$

Equation (3) states that the beliefs are unchanged unless an investment is made in an industry, and equation (4) reflects Bayesian updating of the beliefs about p_i after investing in industry *i* and observing either $y_i(t) = 1$ (success) or $y_i(t) = 0$ (failure).

A. Gittins Index

The Bandit problem is a difficult dynamic programming problem, due to the high dimensionality of the state space. With six industries and beliefs captured by two variables, the resulting state space is twelve dimensional, creating a numerically challenging dynamic programming problem, in particular when this problem must be solved repeatedly as part of an estimation procedure. A breakthrough was made when Gittins and Jones (1974) derived the solution to this problem in terms of the *Gittins Index*. This index is calculated separately for each industry, and the *index result* shows that the optimal strategy is to choose the industry with the highest value of the index. Let $v_i(t)$ be the index for industry i, at time t, and the optimal strategy is

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

The Gittins index is central for the analysis below. While there is no (known) closed form solution for the index, it can be calculated numerically by solving a reduced dynamic programming problem for each industry separately (see Gittins (1989)). To understand the properties of the index, it is helpful to review an approximation, derived by Gittins and Jones (1979), on the form

$$v(a_i, b_i) = \lambda_i + Option\ Value(\lambda_i, n_i)\,, \tag{6}$$

where *Option Value* is a non-negative tabulated function. The net present value of an investment is given by the index, and this value is now decomposed into two terms. The first term, λ_i, is the expected immediate return. This is the value of the investment without any learning, and clearly the total value is at least this large. The second term is the value of the investment in excess of the immediate return, and this represents the value of information or the option value of learning. This value is illustrated in Figure 1 for three levels of n_i. For given λ_i, the option value tends to zero as n_i increases. Intuitively, when the number of investments increases, beliefs become more informed and tighter distributed around λ_i, which in turn approaches the true p_i, and the value of learning vanishes. Conversely, holding n_i fixed, *Option Value* is greater for intermediate values of λ_i and vanishes as λ_i approaches zero or one, consistent with the variance of the Beta distribution being greater for intermediate values of λ_i.

*** FIGURE 1 HERE ***

B. Interpretation of Learning

The assumption that investors learn about success probabilities across industries captures several kinds of learning. Investors may be uncertain about their private abilities. For example, the access to deal flow and the ability to screen and work with entrepreneurs are important for VCs. VC differ in these abilities, and new VCs may not fully know their own abilities yet. In addition, VCs may learn about current industry conditions. Historically, VCs have experienced seen numerous cycles where, i.e. hardware, web-based software, and biotechnology have come in and out of fashion, and the investors learn about these cycles from their current investments.

Finally, the model captures learning at the finer levels of individual ideas or technologies, aggregated to the industry level. New technologies and business models are fraught with uncertainty, and learning about the viability of those is closely related to Arrow (1969)'s definition of innovation, stating that "[t]echnological progress is in the first instance the reduction in uncertainty." The option to make follow-up investments in other companies with similar technologies may be a substantial part of the value of an investment in an entrepreneurial company and a driver of technological progress. Two problems prevent the model from explicitly capturing learning at the level of individual technologies or ideas. The estimation requires a fairly large number of observations of investments in each category, and defining categories at a finer level reduces these numbers. Further, the assumptions that the environment is stationary and that investments are only informative about other investments in the same category become more problematic at finer levels. Without these assumptions, the index result fails, complicating the analysis.

II. Empirical Implementation

A. Description of Sample of VC Investments

The data are provided by Sand Hill Econometrics (SHE) and contain the majority of VC investments in the U.S. in the period 1987 to 2005.[7] SHE combines and extends

[7] It may be a concern that only few companies go public after 2000. For robustness, the model is estimated restricting the sample to end in 2000, 1998, 1996, 1994, and 1992. The main results are robust across these sub periods. The signs and economic magnitudes of the main coefficients (unreported) are largely unchanged, and although the statistical significance is reduced with the smaller sample size, the main coefficients remain statistically significant.

two commercially available databases, Venture Xpert (formerly Venture Economics) and VentureOne. These are extensively used in the VC literature (i.e. Kaplan and Schoar (2005) and Lerner (1995)), and Gompers and Lerner (1999) and Kaplan, Sensoy and Strömberg (2002) investigate the completeness of Venture Xpert and find that it contains most investments and that missing investments tend to be less significant ones.

The sample is restricted to investments made before 2000, since it typically takes companies three to five years after the initial investment to go public or be acquired, and information about these outcomes is current as of 2005. It is common for multiple VCs to invest in the same company, and the sample contains these multiple investments. VCs also typically stage their investments, but the sample is restricted to each VC's initial investment in a company, to focus the analysis on learning from individual companies and the effect on subsequent investments in other companies. While it would be interesting to study learning from individual rounds, the absence of round level outcome measures prevents this. Further, VCs that make less than 40 investments in the full sample are excluded, since their short investment histories make it difficult to draw inference about their learning process and this create convergence problems for the estimation procedure. This reduces the sample from 3,364 to 216 VCs and eliminates around 50% of the companies. Not surprisingly, eliminated VCs are smaller and more idiosyncratic with lower success rates than the remaining ones. The average success rate for the investors in the final sample is 50% (see Table I below). The rate for eliminated investors is only 39%. Overall, the final sample contains 19,166 investments in 6,076 companies by 216 VC firms.

B. Company Characteristics

Each company is classified as belonging to one of six industries: "Health / Biotechnology," "Communications / Media," "Computer Hardware / Electronics," "Software," "Consumer / Retail," and "Other."[8] The distribution of investments across industries is presented in Table I. The classifications are aggregated from twenty-five minor classifications, and it is necessary to ensure that investors have sufficiently long investment histories within each industry classification. The classifications are aggregated such that experience in one industry is more informative about subsequent investments in this same industry but not across industries, as assumed in the model.

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

For each investment, the outcome is given by the binary variable *Success*, and *Success Rate* measures each investor's performance as the number of past successful investments divided by the total number of past investments. An investment is successful when the company eventually goes public or is acquired. This classification is consistent with VCs generating most of their returns from a few successful investments, and is common in the VC literature. Ideally, success would be measured in dollars or as a percentage return, but these data are not widely available. To address concerns about the robustness of this binary outcome measure, Gompers and Lerner (2000) compare

[8] Since learning may be less pronounced for investments in the "Other" category, the model is also estimated while excluding companies in this category. The empirical results (unreported) are unchanged.

different measures, including counting acquisitions as unsuccessful, and find that the measures are highly correlated and lead to qualitatively similar results. In Table I, panel B, the average success rate is 50.3%, ranging from 13.3 to 86.4% across the VCs in the sample.

Finally, the companies are classified as either early-stage or late-stage when they are funded. Late-stage roughly corresponds to the company having regular revenues, and the binary variable *Stage* equals one for these companies; 28.7% of the investments are in such companies.

C. Market Conditions

In addition to learning from their own investments, investors may be affected by public market signals and general market conditions. Two variables are used to control for these effects. *Industry Investments* measures the total number of VC investments in each industry in each year. It varies from 36 investments in "Other" in 1994 to 3,443 investments in "Computer Hardware / Electronics" in 2000. Alternatively, following Gompers, Kovner, Lerner and Scharfstein (2005), the variable *Industry IPOs* contains the number of VC-backed companies going public in each industry for each year. Industries with more IPOs may represent more profitable investment opportunities and attract more VCs, and Gompers, Kovner, Lerner and Scharfstein (2005) find that this is an important determinant of investments. The results confirm their finding.

D. Calculating Expected Returns and Option Values

For each investment the investor's expected immediate return and the option value of learning are calculated from the updated beliefs. Initial beliefs are assumed to be distributed $Beta(a_{i,0}, b_{i,0})$ with $a_{i,0} = 1$ and $b_{i,0} = 19$, and this particular choice is discussed in detail below. The updated beliefs are calculated by counting the number of previous successful and unsuccessful investments, as discussed above, and the resulting counts are denoted by a_i and b_i, respectively. The expected immediate return is calculated as $\lambda_i = a_i / (a_i + b_i)$, and in Table I this return is found to equal 26.0% on average, varying between 3.2 and 71.7%.

The option value is calculated by subtracting the expected return from the Gittins index. This index is calculated using a numerical algorithm from Gittins (1989) with a discount factor of $\delta = 0.99$.[9] With an average time between investments of 48 days, this factor corresponds to an annual discount rate of 8% (the results are robust across a wide range of discount factors). As reported in Table I, *Option Value* is found to equal 6.2% on average, ranging from 1.9 to 8.2%. As a fraction of total value, option value varies from 3.1 to 53.1% with an average of 25.2%.

One concern is that the updating of the investors' beliefs assumes that the investors immediately learn whether an investment will be successful or not. Clearly, this is a stark assumption. However, the opposite assumption, that investors must wait

[9] A MatLab program for this calculation is available from the author.

until the company has an exit event before learning anything, is equally stark. In reality, investors learn gradually during their involvement with a company, but this gradual resolution of the uncertainty is not the focus of the model, and the empirical implementation requires the specification of a point where there learning takes place. Arguably, this is somewhat arbitrary, but a natural point of learning is the point of the initial investment, since this point is well defined for all investments, regardless of their duration or outcome. To the extent that learning takes place later, this introduces an element of measurement error in the investors' beliefs. To investigate this concern, the model is estimated (unreported) where the updating is delayed from 30 to 360 days after the time of the initial investment. Consistent with hypothesis that the beliefs with delayed updating contain less measurement error, the magnitudes of the estimated coefficients are weakly increasing in the delay before the beliefs are updated. However, all estimated coefficients are statistically significant and qualitatively similar.

E. Econometric Specification

VCs' investment decision across industries is specified as a multinomial discrete choice model. At the time of each investment, the value of an investment in industry *i* is

$$v_i = \lambda_i \beta_1 + Option\ Value_i \beta_2 + X_i' \beta_3 + \varepsilon_i \,. \tag{7}$$

The VC invests in the industry with the highest value, and the probability that of choosing industry *i* is denoted π_i. When ε_i follows an *i.i.d.* extreme value distribution, the model is equivalent to a Multinomial Logit model (see McFadden (1973) and McFadden (1974)), and it is well known that

$$\pi_i = \frac{\exp(v_i)}{\sum_{k=1,\dots K} \exp(v_k)} . \quad (8)$$

The scale of the coefficients is not identified, and it is normalized by fixing the variance of the error term to equal one. Note that a general "investor quality" is not identified either. An investor fixed effect entering additively across industries cancels out of the maximum operator. This means that the model cannot measure an investor fixed effect, but the estimates are consistent with the presence of such an effect, even if its changes over time (i.e. if investors learn about or improve their "general quality" in addition to their industry specific p_i).

A high value of β_1 means that investors are more likely to invest in industries with higher expected immediate returns (higher λ). In other words, β_1 captures investors' tendency to *exploit*, and the Bandit model predicts that this coefficient is positive. A high value of β_2 indicates that investors place weight on the option value and invest in companies with a greater value of information, corresponding to more explorative behavior. Finally, the model predicts that the optimal trade-off is given by $\beta_1 = \beta_2$.

III. Evidence of Learning

Estimates of several specifications of equation (7) are reported in Table II. Across all specifications the coefficients on both λ_i and $\textit{Option Value}_i$ are positive and significant. Not surprisingly, investors *exploit* and prefer industries with a higher expected immediate return (i.e. $\beta_1 > 0$), but investors also *explore* and invest in industries with a greater value of learning (i.e. $\beta_2 > 0$).

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

Specifications 2 to 5 control for other factors that may affect investment decisions. Although not part of the formal model, investors may learn from other investors and from public market signals. Specifications 2 and 3 control for market-wide learning or trend by including the total number of VC-backed IPOs (*Industry IPOs*) and the total number of VC investments (*Industry Investments*) during each year in each of the six industries. These market conditions have small but positive and significant effects on investment decisions, consistent with Gompers, Kovner, Lerner and Scharfstein (2005) who document that VCs follow public market signals. However, including these additional controls does not eliminate the effects of λ and *Option Value*, and after controlling for the general trends in the market, investors still appear to internalize the option value of learning.

Specification 4 includes the investors' experience in individual industries (*Industry Experience*), calculated as the total number of past investments the VC has made in each industry. The estimated coefficient is positive and significant, and this variable captures other factors that may lead VCs to concentrate investments in industries where they have longer investment histories and more experience. VCs may be specialized or entrench themselves in an industry, perhaps to enjoy wider access to the deal flow and excluding entrants. One mechanism is suggested by Hochberg, Ljungqvist and Lu (2007) who find evidence that VC syndication networks facilitate such barriers to entry.

Specification 5 is a kitchen-sink regression that includes all the regressors and industry fixed effects. The variable *Previous* is a binary variable that equals one for the industry of the investor's previous investment. This captures additional persistence in investment decisions. For example, if investors only partly update their beliefs between investments, they would be more likely to invest in the same industry repeatedly, and the coefficient on *Previous* would be positive (conversely, if they diversify their investments across industries before updating their beliefs, it would be negative). Overall, the positive and significant coefficients on *Industry Experience* and *Previous* reveal some degree of persistence unexplained by the model. However, controlling for these effects, along with *Industry IPOs* and *Industry Investments*, the coefficients on λ and *Option Value* remain positive and significant, confirming that learning is a significant determinant of investment decisions.

To provide further evidence for the model, Table III presents estimates of a Probit model where the outcome of each investment is a function of investor and market characteristics. Specification 1 is the baseline specification. It shows that investments with higher λ, higher *Option Value*, and in late-stage companies are more likely to be successful. In the learning model, λ measures the investor's expected immediate success probability. The positive and significant coefficient reported in Table III confirms that this measure captures the actual success probability, indicating that investors' beliefs in the model capture economic aspects of the outcomes of the actual investments. The model predicts that exploratory investments, which are made mainly for their option value, have lower success rates. An investment that is attractive because λ + *Option Value* is large should only have a success rate of λ, and *Option Value* should be a weaker predictor of success. Note that this is not a formal prediction of the model. If investors underestimate the value of untried industries, investments with higher *Option Value* have better realized outcomes. Still, the coefficient on *Option Value* provides an indication of the fit between the model and the beliefs. Specification 2 includes industry controls and controls for the investor's experience, and the significance of *Option Value* decreases substantially. Finally, specification 3 is a kitchen-sink regression with a number of additional controls for market conditions and investor experience. In this specification the statistical significance of *Option Value* vanishes entirely, but the significance of λ remains largely unchanged, consistent with model.

**** TABLE III ABOUT HERE ***

A. Investment Strategies and Outcomes

The model further predicts that investors that explore more are more successful. To investigate this relationship, the model is estimated separately for each investor. For this estimation, it is convenient to specify the value of an investment as

$$v_i = \left[\lambda_i + OptionValue_i\right] + OptionValue_i \gamma_{j,1} + Industry\,IPOs_j \gamma_{j,2} + \varepsilon_i . \qquad (9)$$

The bracket contains the immediate return plus the option value, i.e. the Gittins index. The second term is option value in excess of the value in the bracket, and investors with positive $\gamma_{j,1}$ exhibit more explorative investment behavior than predicted by the model.

The coefficient $\gamma_{j,2}$ classifies the investment behavior according to how closely it follows general trends in the market, here measured by *Industry IPOs*.[10] A larger value of $\gamma_{j,2}$ corresponds to an investor that follow these trends to a greater extent.

One slightly unusual feature of this specification is that the scale of the equation is normalized by fixing the "coefficient" for the first term (in the bracket) to equal one. This is different from the usual normalization where the standard error of the error term is set to one. When the objective is to compare investment behavior across investors, this normalization has two benefits.[11] The model can now estimate the standard error of the error term in equation (9), providing an estimate of how closely the investors follow the predictions of the model. An investor with a large standard error can be interpreted as making more "opportunistic" or "random" investment decisions. Moreover, the alternative normalization also makes the parameters comparable across investors. Under the traditional normalization, differences in standard errors would be reflected in the scale of the equation, making it difficult to compare coefficients across investors with different tendencies to follow the predictions of this model. .

Technically, the model is estimated by first estimating a standard Logit model and then rescaling the coefficients using the value of the first coefficient. The standard error

[10] The results are largely similar when general trends are measured using *Industry Investments*.
[11] Note that since this is a normalization of the scale of the equation, it is without loss of generality. In particular, it does not imply that the coefficients on the first two terms equal one, in the sense of the first estimated model from equation (7). In fact, the econometric contents of these two models are exactly the same. In the first model, the error term has a standard error equal to one and the first two terms have coefficients of around four to six (from Table II). With the alternative normalization, the first model would have had coefficients equal to one, but the standard error would only be around 1/6 to 1/4.

of the first coefficient provides a measure of how precisely the characteristics of the investor's strategy are estimated, and for investors with shorter investment histories the coefficients are less precise. To adjust, investors are weighted according to the precision of the estimates of their characteristics, with less weight placed on investors with less precise coefficients. One disadvantage of this method is that a small number of investors have negative estimates of the first coefficient of the model, due to random sampling. These are typically investors with short investment histories and imprecisely estimated characteristics. However, these investors appear to have negative values of σ_j, which is difficult to interpret. Since these investors have low weights, all the results are robust to excluding them as well as replacing σ_j with its absolute value. To have interpretable magnitudes of the coefficients, these measures are standardized to have standard deviations equal to one in the sample. Panel B in Table I presents both the scaled and raw estimates.

Treating each investor as an individual observation, and using *Success Rate* as the performance measure, estimates of the following regression are reported in Table IV.

$$\textit{Success Rate}_j = \beta_0 + \gamma_{j,1}\beta_1 + \gamma_{j,2}\beta_2 + \sigma_j\beta_3 + \varepsilon_j \qquad (10)$$

A positive estimate of β_1 indicates that investors that explore more by placing more weight on *Option Value* have higher success rates, and this is the prediction of the model. A positive estimate of β_2 indicates that investors that follow general trends more have higher success rates, and the model has no prediction about this parameter. A negative estimate of β_3 indicates that investors that deviate more from the model have lower success rates, and this may be supportive of the model, since the model prescribes the optimal decisions given the beliefs.

**** TABLE IV ABOUT HERE ***

In the first specification in Table IV, panel A, investors with higher $\gamma_{j,1}$ are found to have higher success rates, consistent with the model. An investor that explores more discovers more successful investments and realizes a higher success rate. The greater propensity to explore may be a result of more dispersed prior beliefs or a higher discount factor (a δ closer to one), leading to a higher value of learning. Alternatively, the explorative behavior may be a result of suboptimal investment decisions, but even in this case "excess" exploration should lead to a higher success rate. The coefficient on *Standard Deviation* shows that investors with a higher σ_j have consistently lower success rates, suggesting that investors who deviate more from the learning model or make more "random" or "opportunistic" investments are less successful. The magnitudes of the effects of *Option Value* and *Standard Deviation* are economically meaningful. A one standard deviation increase in exploration (within the sample of investors) is associated with a 2.14 to 2.62% increase in success rate, and a one standard deviation in the "randomness" is associated with a 1.60 to 2.40% drop in success rate. Compared to an average success rate of 50.3% in the sample, these are meaningful effects.

The second specification includes $\gamma_{j,2}$, which has a positive but insignificant coefficient, suggesting that the relationship between investors' performance and their tendency to follow the market is weaker. Specification 3 also includes the total number of investments by the investor (*Final Experience*),[12] and the estimated coefficient is positive but insignificant. Kaplan and Schoar (2005) and Sørensen (2007) find that more experienced VCs make more successful investments, but the evidence here is less conclusive, perhaps due to the right-truncation of the sample, making *Final Experience* a noisy measure of experience.

The outcomes of the individual investments provide a finer view of the learning process. Treating each investment as a separate observation, Panel B in Table IV reports the coefficients for a Probit model where the outcome of each investment is estimated as a function of the investor's strategy and additional controls. The empirical specification is

$$\Pr(Success_i = 1) = \Phi\left(\beta_0 + \gamma_{1,j}\beta_1 + \gamma_{2,j}\beta_2 + \sigma_j\beta_3 + X'_{i,j,t}\beta_4\right). \tag{11}$$

Consistent with the evidence from the investors' success rates, the results show that investments by more explorative investors are more successful, and investments by more "random" or "opportunistic" investors are less successful. The economic

[12] The difference between *Final Experience* and *Total Experience* is that *Final Experience* is calculated once for each investor at the end of the sample. *Total Experience* is calculated at the time of each investment and increases through the sample period.

magnitudes are also similar. A one standard deviation increase in $\gamma_{1,j}$ corresponds to an increase in the success probability from 1.48 to 3.35%, and a similar increase in σ_j corresponds to a decrease in the success probability of 2.12 to 2.99%. The second specification includes the measure of the investor's tendency to follow the market, but this effect is again small and insignificant. The final specification is a kitchen-sink regression with additional controls and fixed effects. Not surprisingly, investments in companies at the late stage are more likely to be successful (15.14%). Investments by more experienced investors are marginally more likely to be successful and investments in industries with more VC-backed IPOs are marginally less successful, which is again consistent with Kaplan and Schoar (2005). Overall, the results at the investment level supports the evidence at the investor level, although the sign on $\gamma_{2,j}$ reverses in the last specification.

B. Investment Speed

The model has implications for the timing of the investments. There are several possible hypotheses. Investors may initially make slow explorative investments, and if these investments are successful, accelerate to benefit from their informational advantage. Alternatively, investors may make quick initial explorations, perhaps to capture first-mover advantages, and then continue at a more measured pace. In the model, the speed of investing is captured by the discount factor. A δ closer to one implies less discounting between investments, equivalent to a greater speed. As a starting point, assume that increasing the speed requires costly effort. This may reflect the cost of searching for new investments or investing in lower quality companies and working harder to improve them. The investor's problem is now

$$V\big(F(t)\big) = \max_{s(t),e} \left[Ey_{s(t)}(t) \mid F(t) \right] - C(e) + \delta(e) E\left[V\big(F(t+1)\big) \mid F(t), s(t) \right]. \qquad (12)$$

Here e is effort, $C(e)$ is an increasing convex cost of faster investing, and $\delta(e)$ is the discount rate, which tends to one as effort increases. This extended model predicts that when the continuation value increases (the last term in equation (12)), the benefit of speed also increases, regardless of whether the continuation value reflects a greater value of learning or a larger immediate return.

This is confirmed empirically. The coefficients of the following OLS regression are reported in Table V.

$$Time_i = \beta_0 + Gittins_i \beta_1 + X_i' \beta_2 + \varepsilon_i. \qquad (13)$$

Time is the number of days since the investor's previous investment, and a longer time is equivalent to a slower speed. To control for investments that are made simultaneously, the sample is restricted to investments that are made at least fourteen days apart.[13] In this

[13] The regression results are similar when all the investments are included, but the hazard model described below has problems estimating the hazard rates for investments that are very close.

sample, the average of *Time* is 80.8 days with a standard deviation of 126. The variable *Gittins* represents the continuation value, given by the investment's Gittins index.[14]

In the first specification in Table V the coefficient on *Gittins* is -119.85. As predicted, investments with greater values are made faster. Specification 2 includes industry and year controls, along with the investor's total experience, and more experienced investors are found to invest faster, although the magnitude of this effect is smaller.

**** TABLE V ABOUT HERE ***

In specifications 3 and 4 the option value and the immediate return enter separately. The model predicts that both coefficients should be negative with similar coefficients, but in specification 3 the sign of the coefficient on *Option Value* is positive. However, in specification 4, when including year and industry controls, both coefficients become negative, although their magnitudes are fairly different. The larger negative coefficient on *Option Value* suggests that investors make explorative investments faster, perhaps to capture first-mover advantages or for other reasons outside the model.

Finally, the investment speed can be captured by a hazard model. In Table V, specifications 5 and 6 report estimates of a Cox hazard model and the results are consistent with the results from the previous specifications. Note, for the hazard model,

---

[14] Formally, the continuation value is the Gittins index scaled by a factor, see Whittle (1982) (p. 214). Notice that a formal solution to this problem would adjust the continuation value to capture the expected future speed and its cost. This problem is not solved here.

coefficients greater than one reflects an increase in the hazard rate, corresponding to a shorter time between the investments (and corresponding to a negative coefficient in the OLS regressions). Again, more valuable investments are made quicker, and this effect is observed for both investments with higher immediate returns and higher option values of learning.

IV. Specification and Interpretation of Prior Beliefs

The investors' prior beliefs are assumed to follow a *Beta*(a_0 , b_0) a distribution, and the initial parameters are estimated as $a_0 = 1$ and $b_0 = 19$. The beliefs are estimated from the restriction that investors place equal weight on immediate returns and option value, and hence that $\beta_1 = \beta_2$. To estimate a_0 and b_0, a grid search is performed over their possible values. For each set of values, the model in equation (7) is estimated, and the restriction $\beta_1 = \beta_2$ is tested. The initial prior is the one with the smallest values of a_0 and b_0 for which this hypothesis is not rejected,[15] corresponding to the most uninformative prior. This procedure is motivated by the inverse relationship between the dispersion of the prior beliefs and the magnitude of the option value (higher values of a_0 and b_0 correspond to more informed and less dispersed beliefs). With more dispersed beliefs, option value is greater, and equation (7) needs to load less heavily on option value to explain the investment history, reducing the estimate of β_2. This gives rise to an

[15] Neither the model with industry specific effect nor the model without these effects is rejected.

inverse relationship between the dispersion of the beliefs and β_2, and together with the restriction $\beta_1 = \beta_2$, it provides a way to estimate the dispersion of the initial beliefs.

This procedure leads to prior beliefs with an initial value of λ_0 of 5%, which is low relative to the empirical success rate. These pessimistic prior beliefs reflect low expected returns from investments in new and untried industries and consequently less exploration and more persistence in industry choice than under more optimistic beliefs. Note that the estimated coefficient on *Option Value* is somewhat sensitive to the choice of prior beliefs, due to its direct relationship to the dispersion of the prior beliefs, as explained above. However, the coefficient on λ is substantially more robust to this choice. Moreover, the pessimistic priors are a robust feature of the model and the low value of λ_0 arises from a number of different specifications, including changing the definition of success to only include IPOs (not acquisitions).

There are natural interpretations consistent with the low value of λ_0. VCs may be specialized and have particular abilities in certain industries. This would explain their low expected success probabilities for investments outside their area of expertise, as reflected in a low λ_0. Alternatively, a low value of λ_0 is consistent with learning taking place at the finer levels of individual ideas or technologies. If investors explore and learn within industries, a failure in with an investment in a certain technology may prompt investors to explore other technologies within the same industry. The model would capture this reluctance to switch industries, even after failed investments, as a low value of λ_0.

Finally, the may be some misspecification of the initial beliefs. The assumption that all investors have identical beliefs across industries is not entirely reasonably, but it is necessary for the empirical implementation. This problem is aggravated by the fact that VCs' histories prior to the sample period are unobserved, leading to the "initial conditions problem" discussed by Erdem and Keane (1996). To investigate the magnitude of this problem, the sample is divided into VCs' early and late investments, and the model is estimated separately for these two sub-samples. Estimates using VCs' initial 10 investments (the early sample) and all later investments (the late sample) are presented in the last two specifications in Table II. For the estimates using the "late sample," the 10 initial investments are used to "burn-in" the VCs' beliefs, and the beliefs should be more accurate in this sub-sample. As observed in Table II, the estimated coefficients are more reasonable for the late sample, whereas the estimates for the initial ten investments are less reasonable, confirming that it is difficult to specify one set of initial beliefs across all investors and industries, but that this problem is smaller after the beliefs have been updated a few times.

V. Summary and Conclusion

This paper demonstrates that VCs learn from past investments and anticipate to learn from future ones. To empirically test for the presence of learning, the paper presents and estimates a learning model based on the Multi-armed Bandit model. In this model the distributions of payoffs from investments are uncertain, but investors learn about these distributions from the outcomes of their past investments. The return from an investment consists of both its immediate return and an option value of learning, and balancing these two sources of value leads VCs to trade-off *exploitation* of investments with larger immediate returns against *exploration* of investments with uncertain but potentially higher returns. For the empirical implementation of this model, it is demonstrated that the statistical *index result* can turn this sequential decision problem into a standard discrete choice problem, substantially reducing the complexity of the estimation procedure. This strategy may be applicable to empirical analysis or learning in other situations.

The empirical results confirm that learning is important for VCs investment decisions, and the two alternative hypotheses, that (1) VCs do not learn, and (2) VCs learn only from past investments and do not internalize the value of future learning are both clearly rejected. VCs exhibit *exploitative* behavior by changing their investments in response to the outcomes of past investments to benefit from higher immediate returns. Further, VCs exhibit *explorative* behavior by directing capital towards new unproven investments and internalizing the option value of the information gained from these investments. The model generates a number of additional predictions that are confirmed empirically. In the cross-section, VCs with more exploratory investment strategies have greater success rates, and VCs making more random investments are less successful. Considering the speed of VC investments, the model predicts that more valuable investments are made quicker, which is confirmed empirically. This last evidence is particularly supportive of the model. The VCs' expectations about immediate returns and option values are calculated from their past investment histories and outcomes entirely independently of the timing of these past investments. The fact that the beliefs calculated in this way predict the subsequent investment speed provides strong evidence that the beliefs calculated under the model correspond to the investors' actual beliefs when making investment decisions.

Bibliography

Amador, Manuel, and Pierre-Olivier Weill, 2006, Learning from Private and Public Observation of Others' Actions, *working paper*.

Arrow, Kenneth, 1962, The Economic Implications of Learning By Doing, *Review of Economic Studies* 29, 155-173.

Arrow, Kenneth, 1969, Classificatory Notes on the Production and Transmission of Technological Knowledge, *American Economic Review, Papers and Proceedings* 59, 29-35.

Bergemann, Dirk, and Ulrich Hege, 1998, Venture capital financing, moral hazard, and Learning, *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, 719-752.

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

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

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

Goldfarb, Brent, David Kirsch, and David Miller, 2007, Was There Too Little Entry During the Dot Com Era? *Journal of Financial Economics, forthcoming*.

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

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

Gompers, Paul, and Josh Lerner, 2000, The Determinants of Corporate Venture Capital Successes: Organizational Structure, Incentives, and Complementarities, in Randall Morck, ed.: *Concentrated Corporate Ownership* (University of Chicago Press, Chicago, IL).

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

Hellmann, Thomas, and Manju Puri, 2000, The Interaction Between Product Market and Financing Strategy: The Role of Venture Capital, *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.

Hochberg, Yael, Alexander Ljungqvist, and Yang Lu, 2006, Venture Capital Networks and Investment Performance, *Journal of Finance, forthcoming*.

Hochberg, Yael, Alexander Ljungqvist, and Yang Lu, 2007, Networking as a Barrier to Entry and the Competitive Supply of Venture Capital, *Working Paper*.

Hochberg, Yael, Alexander Ljungqvist, and Annette Vissing-Jorgensen, 2008, Informational Hold-Up and Performance Persistence in Private Equity`, *Working Paper*.

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

Kaplan, Steven, and Antoinette Schoar, 2005, Private Equity Performance: Returns, Persistence, and Capital Flows, *Journal of Finance* 60, 1791-1823.

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

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

Keller, Godfrey, Sven Rady, and Martin Cripps, 2005, Strategic Experimentation with Exponential Bandits, *Econometrica* 73, 39-68.

Kortum, Samuel, and Josh Lerner, 2000, Assessing the contribution of Venture Capital to Innovation, *RAND Journal of Economics* 31, 674-692.

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

Linnainmaa, Juhani, 2006, Learning from Experience, *working paper, University of Chicago*.

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

March, Jim, 1991, Exploration and Exploitation in Organizational Learning, *Organization Science* 2, 71-87.

McFadden, Daniel, 1973, Conditional Logit Analysis of Qualitative Choice Behavior, in Paul Zarembka, ed.: *Frontiers in Econometrics* (Academic Press, New York).

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

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

Morris, Stephen, and Hyun Song Shin, 2002, The social value of public information, *American Economic Review* 92, 1521-1534.

Pastor, Lubos, Lucian Taylor, and Pietro Veronesi, 2006, Entrepreneurial Learning, the IPO Decision, and the Post-IPO Drop in Firm Profitability, *working paper, University of Chicago*.

Quindlen, Ruthann, 2000. *Confessions of a Venture Capitalist: Inside the High-Stakes World of Start-up Financing* (Warner Books, New York).

Rothschild, Michael, 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.

Sørensen, Morten, 2007, How Smart is Smart Money? A Two-Sided Matching Model of Venture Capital, *Journal of Finance, forthcoming*.

Weitzman, Martin, 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).

Figure I:

The figure illustrates Option Value as a function of Lambda, keeping the number of investments constant. Option values are plotted for N equal to 40, 30, and 20, respectively.

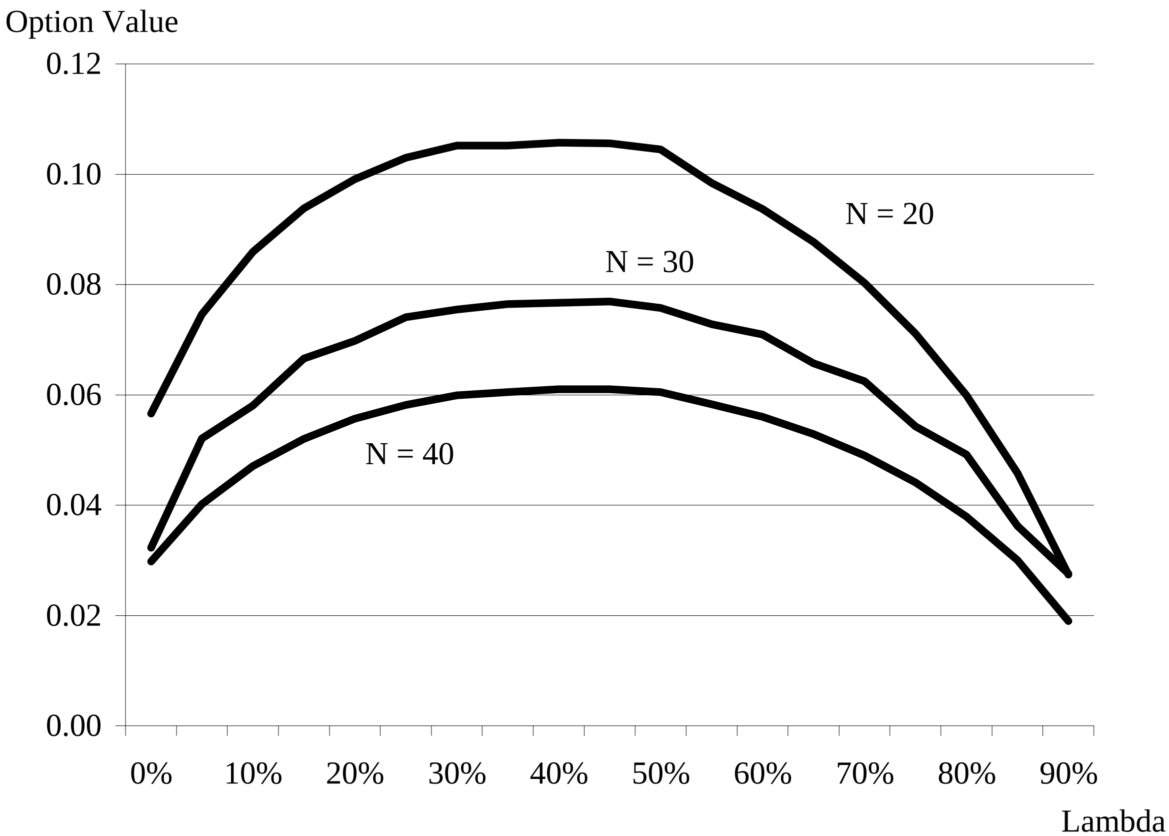

TABLE I: Summary Statistics

The table presents summary statistics of the sample. *IPO*, *Acquisition*, and *Success* are the fraction of companies going public, being acquired, and being classified as successful (IPO + Acquisition), respectively. *Year* is the year of the investment. Panel A presents characteristics at the company level. In Panel B characteristics are presented at the VC level, and the classifications of investment strategies are estimated as described in the text (normalized classifications have std. dev. equal to one in the sample). Panel C presents investment level characteristics. Values of *Lambda*, *Option Value*, *Gittins Index*, and *Option Value / Gittins Index* are calculated for the investments actually made. Across all potential choices, the means of these variables are 0.208, 0.062, 0.270, and 0.311, respectively. Panel D presents investments and IPOs per industry per year for the entire sample, included investors making less than 40 investments.

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 Deviation	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 Deviation	216	0.480	1.000	-6.933	7.634
Industry IPOs	216	0.115	1.000	-11.040	3.602

TABLE I: Summary Statistics (cont.)

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

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)	(0)	(0)	(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	1,025
	(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	1,376
	(40)	(17)	(30)	(5)	(33)	(6)	(131)
1996	472	429	291	176	458	93	1,919
	(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	1,938	1,055	1714	1,249	175	6,975
	(14)	(94)	(27)	(53)	(68)	(3)	(259)
2000	961	2,824	3,443	1,388	1,034	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: 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 *Option Value* 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. Early Sample and Late Sample are estimated using each investors initial 10 investments, and investments 11 and up, respectively. 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	Early Sample	Late Sample
Lambda	4.7657 ***	4.2445 ***	4.3202 ***	3.3410 ***	2.4023 ***	17.8877 ***	4.6238 ***
	(.0718)	(.0788)	(.0780)	(.1354)	(.1484)	(1.7681)	(.1008)
Option Value	6.4545 ***	5.3043 ***	3.0443 ***	17.2952 ***	8.5117 ***	-33.0524 ***	3.7897 ***
	(.8884)	(.9006)	(.9344)	(1.2284)	(1.3057)	(8.7047)	(1.0749)
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	No	No
Observations	19,166	19,166	19,166	19,166	19,166	2,160	17,006

TABLE III: Investment Outcomes

The table reports marginal effect from estimates of a Probit model where the outcome (success or failure) of each investment is the endogenous variable. *Lambda* and *Option Value* 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 IPOs* is the number of companies in the same industry going public in the year of the investment. *Industry Experience* is the past number of investments by the investor in the industry and *Total Experience* is total number of the investor's past investments across all industries. In the third specification, investors initial ten investments are discarded from burn-in. 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	
Lambda	0.3631	***	0.3544	***	0.6668	***
	(0.0368)		(0.0541)		(0.0772)	
Option Value	3.1726	***	1.0309	*	0.5607	
	(0.3975)		(0.5772)		(0.6615)	
Stage	0.1656	***	0.1601	***	0.1726	***
	(0.0141)		(0.0142)		(0.0147)	
Industry Experience			-0.0032	***	-0.0037	***
			(0.0008)		(0.0098)	
Total Experience			0.0004	***	0.0008	***
			(0.0001)		(0.0002)	
Industry IPOs					-0.0004	
					(0.0004)	
Industry Investments					0.0000	
					(0.0000)	
log(Industry Experience +1)					-0.0363	**
					(0.0159)	
log(Total Experience +1)					-0.0397	***
					(0.0159)	
Year Controls	Yes		Yes		Yes	
Industry Controls	No		Yes		Yes	
Observations	19,166		19,166		17,006	

TABLE IV: 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). *Total Experience* measures the number of previous investments by the investor at the time of each investment. *Final Experience* is the 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 Deviation	-0.0160	(.0022)	***	-0.0212	(.0071)	***	-0.0240	(.0080)	***
Industry IPOs				0.0085	(.0112)		0.0086	(.0117)	
Final 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 Deviation	-0.0299	(.0062)	***	-0.0242	(.0075)	***	-0.0212	(.0076)	***
Industry IPOs				0.0029	(.0094)		-0.0013	(.0095)	
Stage							0.1514	(.0157)	***
Total 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 V: Investment Speed

The table reports estimated coefficients from four OLS regressions and two Cox Hazard models. An observation is an investment, and the time since the previous investment (measured in days) is the endogenous variable. Each investor's initial investment and investments made less than 14 days apart are excluded. *Gittins* is the Gittins index of the investment, *Option Value* is the option value, and *Lambda* is the expected immediate return. *Total Experience* is the investor's experience, measured as the total 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.

	OLS								Cox Hazard			
	1		2		3		4		5		6	
Gittins	-119.85	***	-119.54	***					3.88	***		
	(9.00)		(13.06)						(0.28)			
Option Value					385.70	***	-822.23	***			34.23	***
					(118.47)		(149.28)				(36.86)	
Lambda					-108.55	***	-116.93	***			2.35	***
					(10.30)		(12.87)				(0.26)	
Total Experience			-0.35	***			-0.45	***			1.01	***
			(0.03)				(0.04)				(0.00)	
Constant	118.12	***	128.80	***	82.75	***	174.50	***				
	(3.68)		(5.84)		(9.90)		(12.47)					
Industry Controls	No		Yes		No		Yes		No		Yes	
Year Controls	No		Yes		No		Yes		No		Yes	
Observations	10,881		10,881		10,881		10,881		10,881		10,881	