Title
The Risk and Return of Venture Capital

Author
Cochrane, John

Publication Date
2000-01-04

The Risk and Return of Venture Capital

John H. Cochrane[1]

January 4, 2001

[1]Graduate School of Business, University of Chicago. On leave 2000-2001 to Anderson Graduate School of Management, UCLA, 110 Westwood Plaza, Los Angeles CA 90095-1481, john.cochrane@anderson.ucla.edu. This paper is an outgrowth of a project commissioned by OffRoad Capital. I am grateful to Susan Woodward of OffRoad Capital, who suggested the idea of a selection-bias correction for venture capital returns, and who also made many useful comments and suggestions. I gratefully acknowledge the contribution of Shawn Blosser, who assembled all the data used in this paper. Revised versions of this paper can be found at http://gsbwww.uchicago.edu/fac/john.cochrane/research/Papers/.

Abstract

This paper measures the mean, standard deviation, alpha and beta of venture capital investments, using a maximum likelihood estimate that corrects for selection bias. Since firms go public when they have achieved a good return, estimates that do not correct for selection bias are optimistic.

The selection bias correction neatly accounts for log returns. Without a selection bias correction, I find a mean log return of about 100% and a log CAPM intercept of about 90%. With the selection bias correction, I find a mean log return of about 5% with a -2% intercept. However, returns are very volatile, with standard deviation near 100%. Therefore, arithmetic average returns and intercepts are much higher than geometric averages. The selection bias correction attenuates but does not eliminate high arithmetic average returns. Without a selection bias correction, I find an arithmetic average return of around 700% and a CAPM alpha of nearly 500%. With the selection bias correction, I find arithmetic average returns of about 57% and CAPM alpha of about 45%.

Second, third, and fourth rounds of financing are less risky. They have progressively lower volatility, and therefore lower arithmetic average returns. The betas of successive rounds also decline dramatically from near 1 for the first round to near zero for fourth rounds.

The maximum likelihood estimate matches many features of the data, in particular the pattern of IPO and exit as a function of project age, and the fact that return distributions are stable across horizons.

1 Introduction

This paper analyzes the risk and return of venture capital investments. My objective is to measure the expected return, standard deviation, alpha, beta and residual standard deviation of venture capital investment projects.

I use the VentureOne database. The typical data point gives the investment at each round of financing and number of shares. If the firm is acquired, goes public, or goes out of business, we can then compute a return for the venture capital investor. These returns are the basic input to the analysis.

Overcoming *selection bias* is the central hurdle in evaluating venture capital investments, and it is the focus of this paper. Most importantly, firms go public when they have experienced a good return, and many firms in the sample remain private. Therefore, the return to ipo, measuring only the winners, is an upward biased measure of the ex-ante returns to a potential investor.

I overcome this bias with a maximum-likelihood estimate that identifies and measures the increasing probability of going public or being acquired as value increases, the point at which firms go out of business, and the mean, variance, alpha and beta of the underlying returns. The model captures many of the surprising features of the data, such as the fact that the return distribution is little affected by the time to ipo. The estimate also corrects for additional selection biases due to data errors. For example, I am only able to calculate a return for 3/4 of the ipos and 1/4 of the acquisitions, due to data problems. Simply throwing these presumably successful firms out of the sample would bias the results.

I use only returns from investment to ipo or acquisition, or the information that the firm remains private or has gone out of business. I do not attempt to fill in valuations at intermediate dates. There are no data on market values of venture capital projects between investment and exit, so such an imputation requires assumptions and proxies. I also do not base the analysis on returns computed between financing rounds. Though each financing round does establish a valuation, and such returns are potentially interesting, venture capital investors typically cannot take any money out at intermediate financing rounds; they must hold investments all the way to ipo, acquisition or failure. I compute returns to venture capital *projects*. Since venture funds often take 2-3% annual fees and 20-30% of profits at ipo, returns to investors in venture capital funds are often lower.

Results

I verify large and volatile returns *if* there is an ipo or acquisition, i.e. if we do not account for selection bias. The average return to ipo or acquisition is an astounding 698% with a standard deviation of 3,282%. The distribution is highly skewed; there are a few truly outstanding returns of thousands of percent and many more modest returns of "only" 100% or so. I find that returns to ipo/acquisition are very well described by a lognormal distribution. The average log return to ipo or acquisition is still enormous with a 108% mean and a 135% standard deviation. Interestingly, these total returns are quite stable across horizons, and annualized returns are not stable across horizons. As I will explain, this is the pattern we expect of a selected sample. A CAPM in levels gives an alpha of 462%; a CAPM in logs still gives an astonishing alpha of 92%.

The estimates of the underlying return process with a selection bias correction are much more modest and sensible. The estimated average log return is 5.2% per year. A CAPM in logs gives a beta near one and a slightly negative intercept. However, I find arithmetic average returns of 57% and an arithmetic CAPM intercepts or alphas of around 45%. Though these are large, they are still less dramatic than the 698% average return or 462% alpha I obtain without a sample selection correction.

The difference between logs and levels results from the large standard deviation of these individual firm returns, near 100%. This large standard deviation implies an arithmetic average return of 50% or more, even if the average log return is zero. Venture capital investments are like options; they have a small chance of a huge payoff.

Issues

One can cite many reasons why the risk and return of private equity might differ from the risk and return of publicly traded stocks, even holding equal their betas or characteristics such as industry, small size, and financial structure (book/market ratio, etc.)

- Liquidity. Investors may require a higher average return to compensate for the illiquidity of private equity.
- Poor diversification. Private equity has typically been held in large chunks, so each investment may represent a sizeable fraction of the average investors' wealth. Standard asset pricing theory assumes that every investor holds a small part of every risk, and that all assets are held in perfectly diversified portfolios.
- Information and monitoring. Venture capital investments are often not purely financial. The VC investors often provide a "mentoring" or monitoring role to the firm, they sit on boards of directors, and may have the right to appoint or fire managers. Compensation for these activities may result in a higher measured financial return.

On the other hand, venture capital is a competitive business with free entry. If it were a gold mine, we should expect rapid entry. Many venture capital firms are large enough to effectively diversify their portfolios. The special relationship, information and monitoring stories suggesting a restricted supply of venture capital may be overblown. Private equity may be just like public equity.

Literature

Due to the data and econometric hurdles, only a few papers have tried to estimate the risk and return of venture capital. I have found no work that tries to correct for the selection bias.

Long (1999) estimates a standard deviation of 8.23% per year. However, his analysis is based on only 9 unidentified and successful VC investments. Moskowitz and Vissing-Jorgenson (2000) measure returns to all private equity. Venture capital is less than 1% of all private equity, which includes privately held businesses, partnerships, and so forth. They use data from the survey of consumer finances, and use self-reported valuations. They find that a portfolio of all private equity has a mean and standard deviation of return very close to that of the value weighted index of publicly traded stocks.

A natural way to estimate venture capital returns is to examine the returns of venture capital funds, rather than the underlying projects. This is not easy either. Most venture capital funds are organized as limited partnerships rather than as continuously traded or even quoted entities. Thus, one must either deal with missing data during the interim between investments and payout, or somehow mark the unfinished investments to market. Bygrave and Tymmons (1992) found an average internal rate of return of 13.5% for 1974-1989. The technique does not allow any risk calculations. Venture Economics (2000) reports a 25.2% 5 year return and 18.7% 10 year return for all venture capital funds in their data base as of 12/21/99, a period with much higher stock returns. This calculation uses year-end values reported by the funds themselves.

Gompers and Lerner (1997) measure risk and return by periodically marking to market the investments of a single private equity group. They find an arithmetic average annual return of 30.5% (gross of fees) from 1972-1997. Without marking to market, they find a beta of 1.08 on the market. Marking to market, they find a higher beta of 1.4 on the market, and 1.27 on the market with 0.75 on the small firm portfolio and 0.02 on the value portfolio in a Fama-French three factor regression. Clearly, marking to market rather than using self-reported values has a large impact on risk measures, though using market data to evaluate intermediate values almost mechanically raises betas. They do not report a standard deviation, though one can infer from $\beta = 1.4$, $R^2 = 0.49$ a standard deviation of $1.4 \times 16/\sqrt{0.49} = 32\%$. (This is for a fund, not the individual projects.) Gompers and Lerner find an intercept of 8% per year with either the one-factor or three-factor model, though there is an obvious selection bias in looking at a single, successful group. Reyes (1990) reports betas from 1 to 3.8 for venture capital as a whole, in a sample of 175 mature venture capital funds, however using no correction for selection or missing intermediate data.

Discount rates applied by VC investors might be informative, but the contrast between high discount rates applied by venture capital investors and lower ex-post average returns is an enduring puzzle in the venture capital literature. Smith and Smith (2000) survey a large number of studies that report discount rates of 35 to

50%. However, this puzzle depends on the interpretation of "expected cash flows." If you discount the projected cash flows of a project at 50%, assuming success, but that project really only has a 0.83 (1.25/1.5) chance of success, you have done the same thing as discounting true expected cash flows at a 25% discount rate.

2 Overcoming selection bias

To understand the basic idea for overcoming selection bias, suppose that the underlying value of a venture capital investment grows with a constant mean of 10% per year and a constant standard deviation of 50% per year.

The fact that we only observe a return when the firm goes public is not really a problem. If the probability of going public were independent of the project's value, simple averages would measure the underlying return characteristics. Projects that take two years to go public would have an average return of $2 \times 10\% = 20\%$ and a variance of 2×0.50^2; projects that take 3 years to go public would have an average return of $3 \times 10\% = 30\%$ and a variance of 3×0.50^2 and so forth. Thus, the average of (return/time to ipo) would be an unbiased estimate of the expected annual return and the average of (return^2/time to ipo) would form an unbiased estimate of the variance of annual returns[1].

However, projects *are* much more likely to go public when their value has risen. To understand the effects of this fact, suppose that every project goes public when its value has grown by a factor of 10. Now, every *measured* return is exactly 1,000%. Firms that haven't reached this value stay private. The mean measured return is 1,000% with a standard deviation of zero. These are obviously wildly biased and optimistic estimates of the true mean and risk facing the investor!

In this example, we could identify the parameters of the underlying distribution by measuring the *number* of projects that go public at each horizon. If the true mean return is higher than 10%, or if the standard deviation is higher than 50%, more projects will exceed the 1,000% threshold for going public in the first year. Since mean grows with horizon and standard deviation grows with the square root of horizon, the fractions that go public in one year and two years can together identify the mean and the standard deviation. Observations at many different time periods add more information.

In this example, observed returns tell us nothing about the underlying rate of return, but they do tell us the threshold for going public. The fraction that go public or out of business then tell us the properties of the underlying return distribution.

In reality, the decision to go public is not so absolute. The *probability* of going

[1]This statement applies to log i.i.d. returns. Let r_t denote the log return at time t. Then the two-period log return is r_t+r_{t+1}; its mean is $E(r_t+r_{t+1}) = 2E(r_t)$ and its variance is $\sigma^2(r_t+r_{t+1}) = 2\sigma^2(r_t)$. (i.i.d. implies that there is no covariance term.)

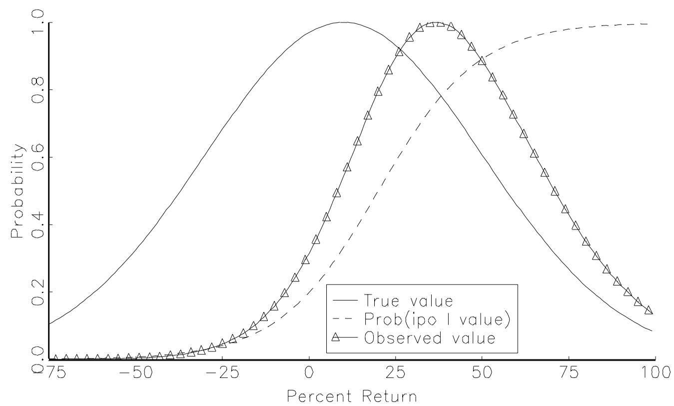

Figure 1: Probability distribution of returns, ("True value"), probability of going public as a function of returns, and observed probability of returns ("Observed value.")

public is an increasing function of the project's value. Figure 1 presents a numerical example that illustrates what happens in this case. The underlying value is normally distributed, graphed in the solid line. The dashed line graphs the probability of an ipo given the return, and rises as the firm's value rises. Multiplying the solid distribution of true values with the dashed probability of going public given value gives the probability of *observing* each return, indicated by the solid line with triangles. While the true mean return is 10%, the mean of the *observed* return is 40%! You can see that the volatility of observed returns is also less than that of true returns, though not zero as it is when all projects go public at the same value.

In this one-period setting, there is really no way to separately identify the underlying value distribution from the probability of going public given value. The "Observed value" line in Figure 1 could have been generated by a true distribution with a 40% mean and a flat probability of ipo given value. However, our data has an extensive time dimension. By watching the shape of such return distributions as a function of the return horizon, and by watching the fraction of firms that go public or out of business at each horizon, we can separately identify the true return distribution from the function that selects firms for ipo.

In a sample without selection bias, the mean and variance of returns keep growing with horizon. In the simple example, the selection-biased return distribution is the same–a point mass at 1,000%–for all horizons. With a smoothly increasing probability of going public, the return distribution will initially increase with horizon, but then will settle down to a constant independent of horizon. This pattern is the signature of a selected sample, and we will see it in the data. This pattern characterizes the economic risks as well. The risk facing a VC investor is as much *when* his return will occur as it is *how much* the return will be.

3 Data

The basic data on venture capital investments come from the VentureOne database. VentureOne collects information on financing rounds that include at least one venture capital firm with $20 million or more in assets under management. I use this data from its beginning in 1987 to June 2000. VentureOne provides the date of the financing round, the amount raised, the post-round valuation of the company, the VC firms involved, and various firm-specific characteristics (industry, location, etc.). They also include a notation of whether the company has gone public, been acquired, or gone out of business, with the associated date of any such event.

VentureOne claims that their database is the most complete source for this type of data, and that they have captured approximately 98% of such financing rounds for 1992 through the present. Therefore, the VentureOne database mitigates a large source of potential selection bias in this kind of study, the bias induced by only looking at successful projects. However, the VentureOne data is not completely free of survival bias. They sometimes search back to find the results of previous rounds, (rounds that did not involve a VC firm with $20 million or more in assets). Gompers and Lerner (2000, p.288 ff.) discuss this and other potential selection biases in the database.

The VentureOne data does not include the financial result of a public offering, merger or acquisition. To compute such values, we used the SDC Platinum service Corporate New Issues databases. We used this database to research the amount raised at ipo and the market capitalization for the firm at the offering price. For companies marked as ipo by VentureOne but not on the SDC database, we used MarketGuide and other online resources.

To compute a return for acquired firms, we used the SDC platinum service Mergers and Acquisitions (M&A) database. We found the total value of the consideration received by companies that underwent a merger or acquisition, according to the VentureOne database. In some instances, even if the company was matched to the SDC M&A database, no valuation information was available for the consideration received. Transactions involving private companies are less likely to be reported to the public.

Using these sources, the basic data consist of the date of each investment, dollars invested, and value of the firm after each investment. The VentureOne data do not give the number of shares, so we infer the return to investment by tracking the value of the firm after investment. For example, suppose firm XYZ has a first round that raises $10 million, after which the firm is valued at $20 million. We infer that the VC investors own half of the stock. If the firm later goes public, raising $50 million and valued at $100 million after ipo, we infer that the VC investors' portion of the firm is now worth $25 million – 1/2 of the value of the pre-ipo outstanding stock. We then infer their gross return at $25M/$10M = 250%. We use the same method to assess dilution of initial investors' claims in multiple rounds.

The VentureOne database does not always capture the amount raised in a specific round, and more often the post-round valuation for the firm is missing. In such instances, we are unable to calculate a return for the investors in that round, as well as the return for any investors of prior rounds for the firm. The estimation includes a correction for bias induced by this selection.

4 Characterizing the data

Before proceeding with a formal estimation, I describe the data. I establish the stylized facts that drive the estimation, especially the fraction of rounds that go public, are acquired, or go out of business as a function of age, and the distribution of returns to ipo or acquisition as a function of age. I check that some of the simplifications of the formal estimation are not grossly violated in the data, in particular that the size of projects is not terribly important, and that the pattern of ipo and exit by age is roughly stable over time.

I take a financing *round* as the basic unit of analysis. Each *firm* may have several rounds, and the results of these rounds will obviously be correlated with each other. I discuss this correlation where it affects the results.

4.1 The Fate of VC investments

Table 1 panel A summarizes the data.

A. Basic Statistics

Total number of financing rounds	16852
Number of companies	7765
Average rounds/company	2.17
Percentage rounds with return	31
Total money raised ($M)	114,983

B. Percent of rounds in various exit categories

	Rounds			Money		
	Return data	No return data	Total	Return data	No return data	Total
Ipo	16.5	5.3	21.7	19.7	4.2	24.0
Acquisition	5.7	14.4	20.2	4.4	8.7	13.1
Out of business	8.9	0	8.9	4.6		4.6
Remains private	0	45.5	45.5		50.0	50.0
Ipo Registered	0	3.7	3.7		8.3	8.3

Table 1. Characteristics of the sample. "Return data" denotes the percentage of rounds for which we are able to assign a return; "No return

data" denotes the percentage of rounds for which we are not able to assign a return; for example due to missing or invalid data. The sample extends from January 1987 to June 2000.

We have nearly 17,000 financing rounds in nearly 8,000 companies, representing 114 billion dollars of investments. Table 1 panel B summarizes the fate of venture capital financing rounds. Of 16852 rounds, 21.7% result in an ipo and 20.2% result in acquisition. Unfortunately, we are only able to assign a return to about three quarters of the ipo and one quarter of the acquisitions. Often, the total value numbers are missing or do not make sense (total value after a round less than amount raised), or the dates are missing or do not make sense. 8.9% go out of business, 45.5% remain private and 3.7% have registered for but not completed an ipo. Obviously, we have no returns for these categories.

Weighting by dollars invested can yield a different picture. For example, large deals may be more likely to be successful than small ones, in which case the fraction of dollars invested that result in an ipo would be larger than the fraction of deals that result in an ipo. The "Money" columns of Table 1 panel B show the fate of dollars invested in venture capital. The fraction of dollars that result in ipo is very slightly larger than the fraction of deals, though the fraction of dollars that results in acquisition is slightly lower. Overall, however, there is no strong indication that the size of the investment affects the outcome. This is a fortunate simplification, and justifies lumping all the investments together without size effects in the estimation to follow.

Figure 2 presents the cumulative fraction of rounds in each category as a function of age. By 5 years after the initial investment, about half of the rounds have gone public or been acquired. After this age, the chance of success decreases; more and more rounds go out of business, and the rate of going public or acquisition slows down[2].

One naturally wonders whether age alone is the right variable to track the fortunes of VC investments. Perhaps the fate of VC investments also depends on the *time* that they were started. For example, the late 90s may be a time in which VC investments prosper to ipo unusually quickly. To examine this question, Figure 3 presents the exit probabilities of Figure 2, broken down by date of the initial VC investment.

Figure 3 suggests that things are happening a bit faster now. Any given percentile of firms that go public or are acquired happens about one year sooner in the later subsamples than in the earliest subsample. But this is not just better fortune for VC investments. The fraction that are out of business at any given age has also risen.

[2] The lines in Figure 2 are not exactly monotonic, as cumulative probabilities should be, because the sample is different at each point. For example, the fractions in various states at a 5 year age must be computed for all rounds that start before 1995, while the fractions in various states at a 3 year age is computed for all rounds that start before 1997. Except for the extreme rightmost points, where we can only consider the small number of firms that started in 1987, however, the lines are quite smooth, suggesting that merging rounds with different start dates is not a mistake.

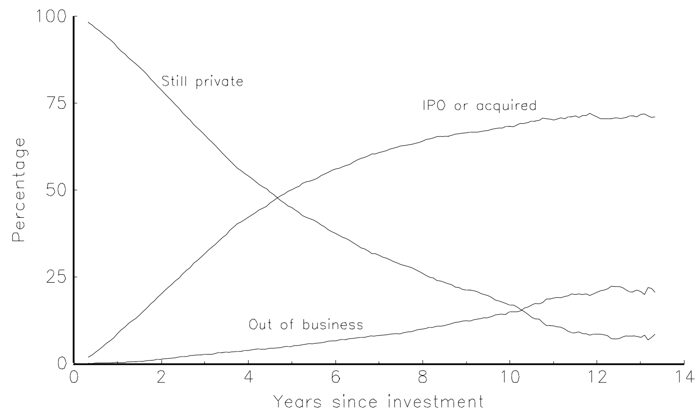

Figure 2: Cumulative exit probabilities as a function of age.

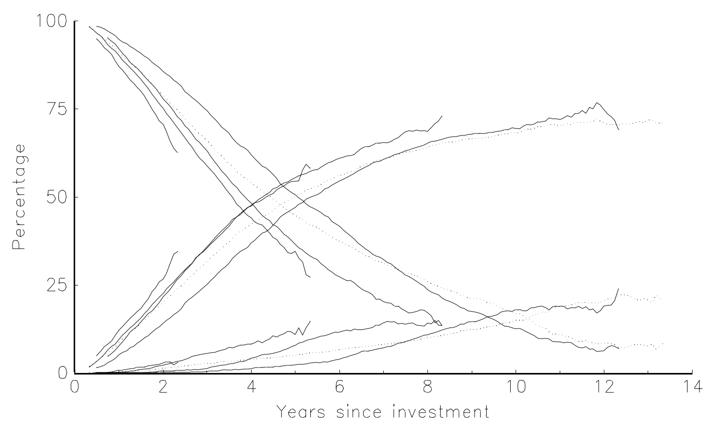

Figure 3: Cumulative exit probabilities as a function of age, for dated subsamples. The subsamples are 1988-1992, 1992-1995, 1995-1998, and 1998-June 2000. In each set of lines, the shorter line is the latest sample date. The longest lines show the full-sample results from Figure 2. The declining lines represent the fraction still private. The upper set of rising lines represent the fraction going public or being acquired. The bottom set of rising lines represent the fraction going out of business.

Despite these differences, however, Figure 3 is reassuring that the overall character of VC investments has not dramatically changed. The basic transition probabilities as a function of age are reasonably stable across the subsamples, and I will use age alone as the state variable in estimation that follows.

4.2 Returns

Our central question, of course, is the return to VC investments. In this section, I characterize what we can see – returns *when there is an ipo or acquisition.* As I emphasized in the introduction, these are *not* the ex-ante returns to VC investments, and they may tell us more about the values that trigger the decision to go public than they do about the underlying rate of return. However, we have to accurately gauge how well things go when they do go well, both for its own interest, and since this is the crucial measurement that I use to calibrate a model that corrects for sample selection.

Net returns

Figure 4 plots a smoothed histogram of the distribution of net returns. (These returns are not annualized; annualized returns follow.) For this and the remaining analysis, I put all investments in together, including multiple rounds in the same company. Thus, round 1 investment to ipo is one return, and round 2 investment to ipo in the same company is another return. The next section includes a separate analysis by round.

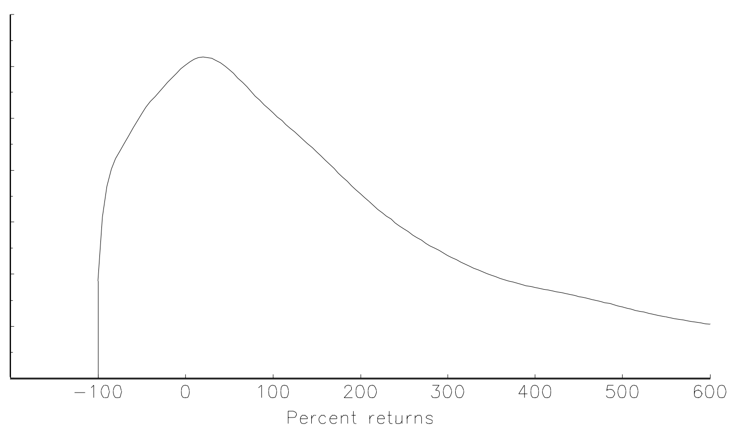

Figure 4: Smoothed histogram (kernel estimate) of the distribution of percentage returns, for firms that are acquired or go public.

Figure 4 shows an extraordinary skewness of returns. Most returns are modest, but there is a long right tail of extraordinarily good returns. 15% of the firms that go public or are acquired give a return greater than 1,000%! It is also interesting how many modest returns there are. About 15% of returns are less than 0, and 35% are less than 100%. An ipo or acquisition is not a guarantee of a huge return. In fact, the modal or "most probable" outcome in Figure 4 is about a 25% return.

	All	0-6 mo.	6m-1yr	1-3yr	3-5yr	5yr+
Number	3595	334	476	1580	807	413
A. Net Returns (percent)						
Average	698	306	399	788	942	535
Std. Dev.	3282	1659	881	3979	3822	1123
Median	184	77	135	196	280	209
25-75 Range	521	225	346	508	719	623
B. Log Returns (percent)						
Average	108	63	93	114	129	97
Std. Dev.	135	105	118	134	144	147
Median	105	57	86	108	133	1.13
75-25 Range	158	114	125	149	172	189

Table 2. Means and variances of returns when there is an ipo or acquisition. Units are percent returns, not annualized.

Table 2 assesses the net return distribution numerically. The first column ("All") of Table 2 summarizes the entire return distribution, corresponding to Figure 4. While the modal return (peak of Figure 4) is near 25%, the median is 184%, and the average return is an impressive 698%. This high average reflects the small possibility of making an astounding return, combined with the much larger probability of a more modest return.

The standard deviation of returns reflects huge volatility and the same skewness. The standard deviation of returns is 3282%. Summing squared returns really emphasizes the few positive outliers! The range between the 25% and 75% quantile, like the median, is a dispersion measure less sensitive to outliers. At 521%, this range is much lower, but still impressive.

Log returns

The skewness of returns suggests a log transformation. Figures 5 and 6 present the cumulative distribution of log returns to ipo or acquisition, and Figure 7 presents the smoothed histogram. The figures include a normal distribution calibrated to the mean and variance of the log return. As you can see, the log transformation does a very good job of capturing the skewness of returns, and the lognormal distribution is a quite good approximation to the distribution of actual returns. The actual distribution has slightly fatter tails than the normal distribution, but the difference is not huge. You can see this most clearly in Figure 6 which blows up the right tail.

The average log return (Table 2) is 108%, nearly equal to its median of 105%. The standard deviation of log returns is a large 135%, while the 25% and 75% quantiles are roughly symmetric about the mean and median. These numbers verify that the log transformation makes the return distribution quite symmetric. But 108% mean and 135% standard deviation are still an extraordinary mean and volatility of returns.

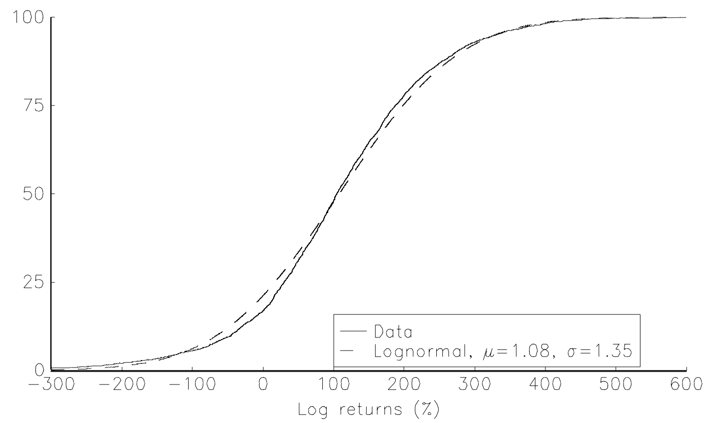

Figure 5: Cumulative distribution of log returns to ipo or acquisition.

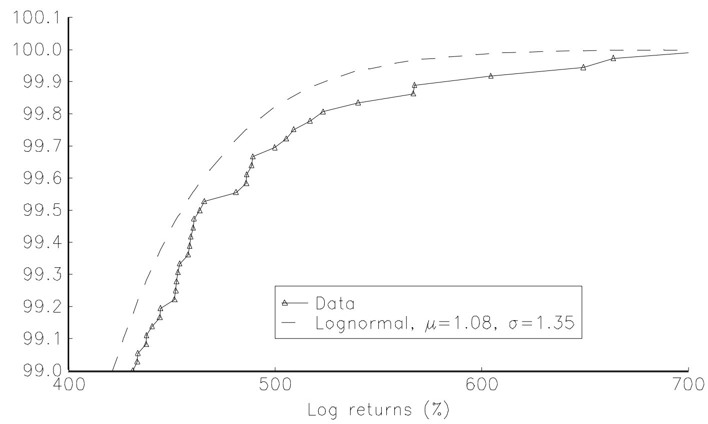

Figure 6: Right tail of the cumulative distribution of log returns, together with a normal distribution fitted to the mean and variance of log returns.

Which kind of return?

From a statistical point of view, it is clearly better to describe moments of the *log* return distribution. However, for portfolio decisions, the expected level or arithmetic average return and the corresponding standard deviation are the important statistics. If you form a portfolio composed of fraction w in a VC investment with return R^{vc} and fraction $1 - w$ in a riskfree return R^f, the return of the portfolio R^P is given by $wR^{VC} + (1 - w)R^f$, with mean $E(R^P) = wE(R^{VC}) + (1 - w)R^f$ and standard deviation $\sigma(R^P) = w\sigma(R^{VC})$. We cannot make this kind of transformation with the mean and variance of log returns. Mean-variance portfolio theory also specifies the actual return rather than the log return. Of course, one can easily transform between the two measures. For example, if the best statistical description is that the log return is normally distributed with mean μ and variance σ^2, then we can compute the actual or arithmetic mean return as $e^{\mu+\frac{1}{2}\sigma^2}$.

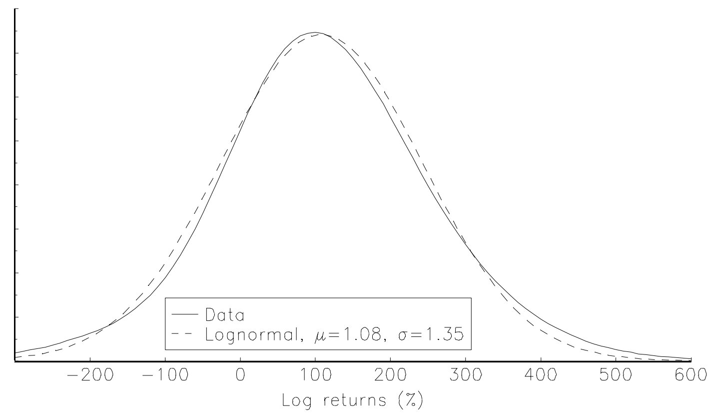

Figure 7: Smoothed histogram of log returns, together with a normal distribution fitted to the mean and variance of log returns.

Returns sorted by age

So far, I have lumped all returns together without consideration of how long it takes to achieve that return. As we will see, this turns out to be a sensible way to characterize the data. However, it is important to understand how returns vary by age of the project. The pattern of returns with age, together with the exit history depending on age, is the central piece of information I use to overcome the selection problem that good projects are much more likely to go public. The remaining columns of Table 2 presents statistics sorted by age, and Figure 8 presents smoothed histograms of log returns sorted by age categories.

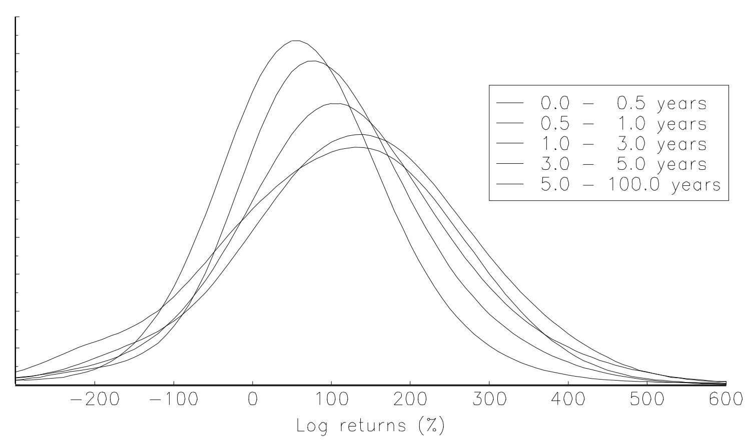

Figure 8: Smoothed histogram of returns to ipo or acquisition, sorted by time between financing and ipo or acquisition. The leftmost and highest curve is the 0-6 month category. The older categories correspond to curves with successively lower peaks.

The distributions in Figure 8 shift slightly to the right, and all the measures of average returns in Table 2 rise as the time to ipo lengthens, up to the 3-5 year category. The 5 year plus curve shifts slightly to the left and the average return in Table 2 decreases. The volatilities also increase slightly as the horizon increases.

However, what is most surprising about both average and volatility is that they do not increase much *faster* with horizon. Stock returns are close to independent over time. Thus, the mean log return should grow linearly with horizon and the standard deviation should grow with the square root of the horizon. Even when they do increase, neither mean nor standard deviation grow anything like this fast.

Instead, the pattern of returns sorted by age in Figure 8 and Table 2 shows the signature of a selected sample. This pattern results if the probability of going public is small and flat below a return of about 200%, but then increases smoothly. With an age below one year, most firms cannot build up the 200% return that it takes to make an ipo likely. Hence, the ipos we see come from the fairly constant and small hazard of ipo in this value region. Projects that take longer to go public have proportionally higher and more volatile returns. As time passes, however, more firms have the time to build up the large values that make in ipo more and more likely. At these return horizons, the return distribution reflects the probability of going public more than it reflects the underlying character of returns. This fact results in return distributions that become stable over different horizons. If a firm achieves a good return in the first year, it goes public; we see the good return and then it is removed from the sample. Most firms in the 5 year return distribution *did not* have a good first year – if they did, they would have gone public. Thus, firms in the 5 year return distribution have a mean less than 5 times that of the firms in the one year distribution.

Betas

Table 3 presents regressions of returns to ipo or acquisition on the S&P500 index return. Again, these provide an interesting baseline and stylized fact, but do not measure the return process of the underlying investments until we correct for selection bias. In fact, the risk facing a VC investor is as much *how long* the project will take to reach ipo, as it is *how large* the eventual return will be, and adverse market movements may in the end contribute more to delay than to value.

The intercepts (alpha) are huge. The beta for net returns is large at 2.04. This is an indication that pre-ipo securities are highly risky, in this conventional sense that they are quite sensitive to market returns. The log returns trim the outliers somewhat, and produce a much lower beta. The R^2 values in these regressions are tiny. Market returns of 10 or 30% are just a tiny fraction of the risks one faces with 700% average returns and 3,000% standard deviations! A scatterplot of these regressions would just be a huge round cloud. This is also why betas are poorly measured. Similar regressions across horizons do not show much consistency or any interesting patterns.

	α	(s.e.)	β	(s.e.)	R^2
All net returns	462	148	2.04	0.83	0.00
All log returns	92	4	0.37	0.07	0.01

Table 3. CAPM regressions, $R_t = \alpha + \beta R_t^m + \varepsilon_t$ and $\ln R_t = \alpha + \beta \ln R_t^m + \varepsilon_t$. Each return to ipo is regressed against the market return for the same period. α is in percent.

Annualized returns

Figure 9 presents the distribution of annualized returns, and Table 4 presents corresponding statistics. Obviously, we usually compare returns at different horizons by annualizing them, so it is natural to try this transformation.

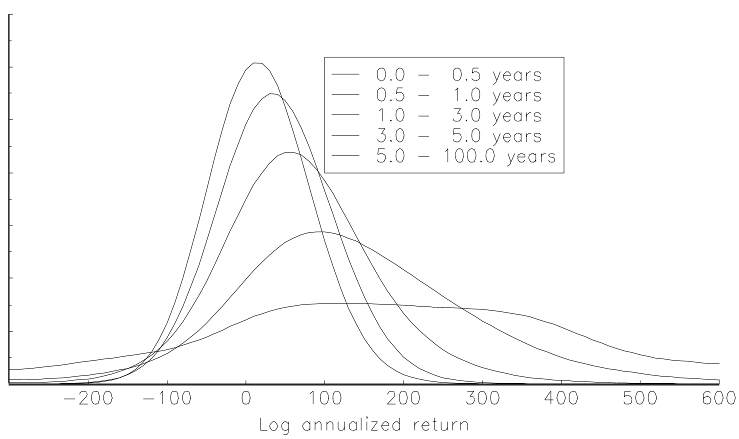

Figure 9: Smoothed histogram of annualized returns, sorted by age category

	All	0-6 mo.	6m-1yr	1-3yr	3-5yr	5yr+
		A. Net Returns (Percent)				
Average	4×10^9	4×10^{10}	2,064	251	54	20
Std. Dev.	2×10^{11}	7×10^{11}	881	3,979	3,822	1,123
Median	62	557	211	79	42	18
25-75 Range	156	3,614	697	152	68	35
		B. Log Returns (Percent)				
Average	72	201	124	64	35	15
Std. Dev.	148	371	165	81	40	24
Median	48	188	114	58	35	17
75-25 Range	86	337	172	80	48	29

Table 4. Means and variances of annualized returns.

The mean and volatility of annualized returns decline sharply with horizon. Comparing annualized returns with actual returns you can see that the *actual* returns are much more stable across age categories than are the *annualized* returns–exactly the opposite of the pattern you should observe for an unselected sample. In an unselected sample of i.i.d. returns, the mean annualized log return should be the same for different horizons. Seeing Table 4, it should now be clear why I summarize the return to ipo data by returns that are *not* annualized.

The annualized return distribution is *extremely* skewed. The mean annualized return is $4\times10^{9}\%$, with standard deviation $2\times10^{11}\%$, though the median and interquartile range are a sensible 62% and 159%. Again, this represents a small probability of a few extremely large returns. Furthermore, the extreme annualized returns result from a sensible return that occurs over a very short time period. If you experience a mild (in this data set) 100% return, but that happens in two weeks, the result is a $100 \times (2^{24} - 1) = 1.67 \times 10^{9}$ percent annualized return. You can see this pattern in the breakout of annualized returns by horizon; the extremes happen all at the short horizons.

Some of these huge annualized returns may result from measurement error in the dates. 3 observations have ipo dates before investment dates, and there are several more with ipo dates one or two months after investment dates. Since they imply such huge annualized returns, I trim all observations with ipo less than two months after financing, and I focus the analysis on the distribution of actual rather than annualized returns, which are less sensitive to measurement errors in the dates.

5 Maximum likelihood estimates

5.1 Maximum likelihood estimation procedure

My objective is to estimate the mean, standard deviation, alpha and beta of venture capital investments, correcting for the selection bias that we do not see returns for projects that remain private. To do this, we have to write a model of the probability structure of the data – how the returns we do see are generated from the underlying value process and the decision to go public or out of business.

Let V_t denote the value of the firm at date t. I model the growth in value as a lognormally distributed variable with parameters μ and σ. These are the central parameters we want to learn about.

$$\ln\left(\frac{V_{t+\Delta}}{V_t}\right) \tilde{} N(\mu\Delta, \sigma^2\Delta). \quad (1)$$

I normalize each project to an initial value of 1. I use a time interval Δ = three months.

Each period, the firm may go out of business, go public, or be acquired. k denotes the lower bound on value. If $V_t \leq k$, the firm goes out of business, for sure:

$$\Pr(\text{out of business}|V_t) = \begin{cases} 0 & V_t > k \\ 1 & V_t \leq k \end{cases} . \tag{2}$$

A lognormal process such as (1) never reaches a value of zero, so we must envision something like k if we are to generate a finite probability of going out of business. I interpret k as a level of leverage or debt. It can also be interpreted as a cutoff rule by investors; they give up and go out of business when value reaches k.

If the firm remains in business, it may go public or be acquired. I do not distinguish the two outcomes in the estimation. The probability of going public is an increasing function of value. I model this probability as a logistic function,

$$\Pr(\text{ipo}|V_t, V_t > k) = 1/(1 + e^{-a(\ln(V)-b)}) \tag{3}$$

This function rises smoothly from 0 to 1 as value increases. (See Figure 10.)

In either of these cases, the firm is removed from the population of firms still in the sample. The probability of being removed is

$$\Pr(\text{removed}|V_t) = \begin{cases} 1/(1 + e^{-a(\ln(V)-b)}) & V_t > k \\ 1 & V_t \leq k \end{cases} . \tag{4}$$

Thus, the probability of having value $V_{t+\Delta}$ at the beginning of period $t + \Delta$ is

$$\Pr(V_{t+\Delta}) = \int dV_t \Pr(V_{t+\Delta}|V_t) \left[1 - \Pr(\text{removed}|V_t)\right] \Pr(V_t) \tag{5}$$

and $\Pr(V_{t+\Delta}|V_t)$ is given by the lognormal distribution of (1).

We do not have valid observations on all out of business firms, since some of the dates are wrong. Thus,

$$\Pr\left(\text{out of business at } t, \text{ see}\right) = c \times \Pr\left(\text{out of business at } t\right). \tag{6}$$

I estimate c directly as the fraction of out of business rounds with valid data. We do not have valid observations on all of the ipo/acquired either. Thus,

$$\Pr(\text{ipo at } t, \text{ value} = V_t, \text{ see}) = d \times \Pr(\text{ipo at } t, \text{ value} = V_t) \tag{7}$$

I estimate d directly as the fraction of ipos and acquisitions with valid data.

Now, for given parameters $\{\mu, \sigma, k, a, b, c, d\}$ I can recursively calculate the probability distribution of values and dates for out-of-business and ipo exits, and the probability of reaching any given age still private with value V_t. I set up a grid of log values, and initialize all probabilities to zero except at value = 1. Using (1), I find the probability of entering period 1 at each value gridpoint. Then, using (2) and (3) and (6) and (7) I find the probability of observing an ipo or bankruptcy in period

1, and the probability of having had in ipo or bankruptcy but generating bad data. Now, using (5), I find the probability of entering period 2 at each point on the value grid, and so on.

Having found the probabilities of all possible events, I loop through the data to compute the likelihood function. The sample consists of observations of venture capital financing round. Each round results in one of the following categories:

1. Ipo/acquired with good data.
2. Ipo/acquired with good dates but bad return data.
3. Ipo/acquired with bad dates and return data.
4. Still private. Age = (end of sample) - (investment date).
5. Out of business, good exit date.
6. Out of business, bad exit date.

Based on the above structure, for given parameters $\{\mu, \sigma, k, a, b, c, d\}$, we can compute the probability of seeing a data point in any one of these categories. Taking the log and adding up this probability over all data points, we obtain the likelihood. For "bad data" observations, I take the corresponding cumulative probabilities. For example, for the second category, I take the probabilities that the firm goes public at date t, we do not see data, and value = V_t, and sum over values. For the third category, I sum over all dates with ages less than (end of sample) - (investment date) as well.

The return to equity if there is an ipo is

$$R_t^e = \max\left(\frac{V_t - k}{1-k}, 0\right) \tag{8}$$

The return to equity if the firm goes out of business is zero (even if $V_t < k$). I use this concept of equity return in the data – when a firm goes public, I use 8 to infer the value of V_t from the returns to shareholders.

Estimates of alpha and beta

To estimate a regression model, I specify

$$\ln\left(\frac{V_{t+\Delta}}{V_t}\right) = \gamma + \ln R_{t+\Delta}^f + \delta(\ln R_{t+\Delta}^m - \ln R_t^f) + \varepsilon_{t+\Delta};\ \varepsilon_{t+\Delta}\tilde{}N(0,\sigma^2) \tag{9}$$

in place of (1). This is like the CAPM, but in log returns rather than levels of returns. I derive the parameters of the CAPM in levels implied by (9) below.

To estimate (9), I group all investments according to the quarter in which they are made. Then, I use the observed time series of R_t^m and R_t^f to find the probability of returns, ipos, out of business etc., for investments that start on that date, i.e. for investments that have that particular experience of market return and interest rate. Each quarter of start date requires a different simulation, so maximizing the likelihood function takes much longer, and requires a coarser value grid. For this reason, I separately report estimates of just mean and standard deviation, and then a smaller number of estimates with alpha and beta as well. (We could avoid a separate simulation for each investment date if the probability of going public depended only on the cumulated *residual* return ε. However, this seems unrealistic – firms seem to go public on the heels of large market return movements as well as after large individual increases in value.)

Comment on identification

You don't get something for nothing, and the apparent ability to separately identify the probability of going public and the parameters of the return process does come by imposing assumptions.

Most importantly, I assume that *the function* $\Pr(\text{ipo}|V_t)$ *is the same for firms of all ages* t. If you double the initial value in a month, you are just as likely to go public as if it takes 10 years to double the initial value. This is surely unrealistic at very short and very long time periods. One might want to have a different function for each age of firm, but then we are basically back to the static problem in which we can only identify the return from the selection probability by functional form assumptions.

I also assume that *the return process is independent over time*. One might specify that value creation starts slowly and then gets faster, or that betas change with size, but identifying these tendencies (and separating them from a direct age effect) will be difficult. The returns on publicly traded firms are far from predictable, so this is probably a less questionable assumption.

The simulation also assumes specific *functional forms* for the return distribution and probability of ipo and bankruptcy. I suspect that this is not a central assumption, but the programs already take so long to run that nonparametric or more loosely parameterized estimates are not feasible.

5.2 Estimates and interpretation

5.2.1 Mean and standard deviation

Table 5 presents maximum likelihood parameter estimates. Table 5 includes standard errors for the mean return μ, which are the only interesting ones. Table 6 includes the remaining standard errors, which are all tiny. I discuss alphas and betas in the next section.

The central parameters are the mean and standard deviation of returns. The mean log return is 5.2%. Compared to mean log returns of 100% or more in the selected sample, accounting for sample selection has a dramatic effect. To put these numbers in perspective, Table 5 includes the mean and standard deviation of the S&P500 index over the sample period. The mean log return of the VC investments was roughly half of the mean log S&P500 return over this sample. (The Jan 87-June 2000 sample covers the entire period of the VentureOne data. However, most of the VC investments are concentrated in the later part of this period, so I present the Jan 91-June 200 S&P index return as a more relevant comparison. This time period was amazingly good for publicly traded stocks.)

The standard deviation of log returns is quite large, 98%; much larger than the roughly 10% standard deviation of the S&P500. These are individual stocks, and so we expect them to be quite volatile compared to a diversified portfolio such as the S&P500. The volatility of individual large publicly traded stocks is typically 50%, and values as high as 93% are not uncommon for small growth NASDAQ stocks. Keep in mind also that the 93% standard deviation is the annualized instantaneous standard deviation. It may be easier to digest as $93/\sqrt{365} = 4.9\%$ per day.

The major effect of the high volatility is to give VC investments a surprisingly high *arithmetic* mean return. The mean arithmetic return is a whopping 56.9%, with standard deviation of 119%. To get the CAPM to explain such a high mean arithmetic return, using a 5% interest rate, we would need $\beta = (56.9 - 5)/(17.6 - 5) = 4.1$

The leverage parameter k is about 5.4%. Whether interpreted as actual leverage, or the decline in value necessary for investors to give up, this is a low number, but reasonable. These firms do not in fact have much debt, and VC investors are likely to hang in there and wait for the final payout.

All of the standard errors are calculated from the second derivatives of the likelihood function. The standard error of the mean return is the only interesting case, as the others are all tiny compared to their estimates. The mean return appears to be quite well measured. However, I have not accounted for any cross-correlation in the returns–this standard error treats each of the 16720 firms as independent observations. This deficiency does not bias the estimates, but it may result in optimistic standard errors. One source of correlation is that there are several rounds in each firm; I address this issue by breaking out the separate rounds below. The most important other element of cross-correlation is likely the dependence on common components, including the market return and returns on other factor portfolios (small stocks, growth/value, industry averages, etc.). The standard error of the mean market return is the subject of an entire literature; reasonable estimates of the equity premium range from 2% to over 10%, even given a hundred years of data. Thus, to the mean return in Table 5, add or subtract your ideas about how the ex-post market return in the last 10 years has diverged from the true mean return; to the standard error add your ideas of the uncertainty in the mean market return. Finally, modeling simplifications are likely a much larger source of uncertainty than econometric details.

	All	S&P500 87-00	S&P500 91-00	Industries Health	Industries Info	Industries Retail	Industries Other	Financing rounds 1	Financing rounds 2	Financing rounds 3	Financing rounds 4
μ	5.17	15.7	17.6	11.0	5.12	-3.03	7.76	7.88	5.09	6.52	2.20
std. err	0.66	3.13	3.13	0.62	0.90	2.51	1.58	0.99	1.20	1.30	2.32
σ	98.0	11.3	9.4	49.8	108	122	45.6	105	98.2	76.9	80.3
$E(R)$	56.9	17.2	19.7	24.1	69.1	77.4	18.6	68.6	57.0	37.7	36.0
$\sigma(R)$	119	13.0	11.3	53.6	137	160	48.4	133	119	87.33	91.2
k (%)	5.4			33.2	3.6	2.6	33.2	3.6	5.4	14.9	12.1
a	0.92			1.02	0.97	0.70	1.03	1.05	1.14	1.14	1.11
b	4.17			3.78	3.96	5.7	3.78	4.24	3.45	3.10	2.82
c	0.95			0.96	0.94	0.96	0.94	0.93	0.97	0.98	0.96
d	0.52			0.54	0.53	0.46	0.27	0.41	0.54	0.63	0.68
N	16720			3917	9232	3129	442	7720	4494	2455	1236

Table 5. Maximum Likelihood estimates. μ and σ describe the underlying growth of value, $\ln\left(\frac{V_{t+1}}{V_t}\right) \tilde{} N(\mu, \sigma^2)$. I report $400 \times \mu$ and $200 \times \sigma$ so the units are annual percentages. $E(R)$ and $\sigma(R)$ describe the corresponding level rather than log return, $E(R) = 400 \times (\exp(\mu + 1/2\sigma^2) - 1)$ and $\sigma(R) = 200 \times \left(E(R)\sqrt{e^{\sigma^2} - 1}\right)$. k is the cutoff value at which the project goes bankrupt, and can be interpreted as the value of debt; the project goes out of business when $V_t \leq k$. a and b describe the probability of going public (or being acquired) as a function of value, $\Pr(\text{ipo}|V_t, V_t > X) = (1 + e^{-a(V-b)})^{-1}$. c is the fraction of out of business firms for which we have good data, and d is the fraction of ipo/acquired firms for which we have good data. "Health" includes Biopharmaceuticals, Healthcare, Medical Devices, Medical IS and Other Medical. "Info" includes Communications, Electronics, Information Services, Other IT, Semiconductors and software. "Retail, svc." includes Consumer Products, Consumer Services, Retailers.

	All	Health	Info	Retail	Other	1	2	3	4
σ	1.0	1.3	1.4	3.5	3.6	1.7	1.7	1.8	3.2
k	0.3	1.8	0.3	0.5	4.4	0.4	0.5	1.2	2.0
a	0.02	0.06	0.02	0.01	0.21	0.03	0.04	0.07	0.09
b	0.08	0.18	0.08	0.08	0.68	0.10	0.11	0.16	0.21

Table 6. Standard errors for estimated parameters other than μ.

Graphical comparison with the data

Figure 10 presents the estimated probability of going public each quarter as a function of value. When the firm has it has increased in value by $e^3 = 20$, the firm has about 30% chance of going public each quarter, or a 61% chance of going public in a year, which seems sensible. However, the function is actually quite flat; the horizontal axis covers the entire range of possible returns. A log value of 6 corresponds to a $100 \times (e^6 - 1) = 40,243\%$ return. This flatness is necessary to generate the wide dispersion we see in returns to ipo. If every firm went public the moment its value increased by 20, then the standard deviation of returns to ipo would be zero, not 3,000%.

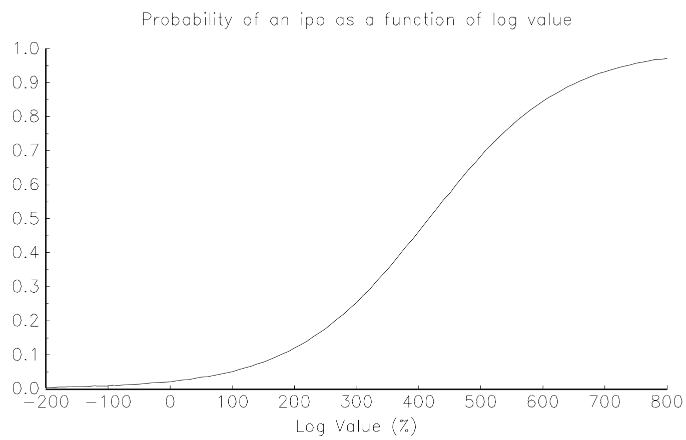

Figure 10: Estimated probability of going public as a function of log value. Log value is expressed in percent units, i.e. 100× log value.

Figure 11 presents the model's predicted probability that a project will wind up in various exit states as a function of its age. Comparing this graph with Figure 2, you can see that the model tracks the corresponding age profiles in the data quite well.

Figure 12 presents the model's predictions for the distribution of log returns at ipo/acquisition, sorted by age category. The most peaked curve is a 3 month horizon. The curves march to the right at 6 month, 9 month, and then 1, 2, 3, ...14 year horizon. You can see that the mean return and standard deviation of return initially increase with horizon, but then the returns settle down to a constant distribution independent of horizon.

There are two offsetting effects. If the probability of going public were not a function of value, then the return distribution would shift to the right and spread out with horizon. This is the dominant effect for short horizons. As you can see in Figure 10, in the value range between -100% and 100%, the probability of going

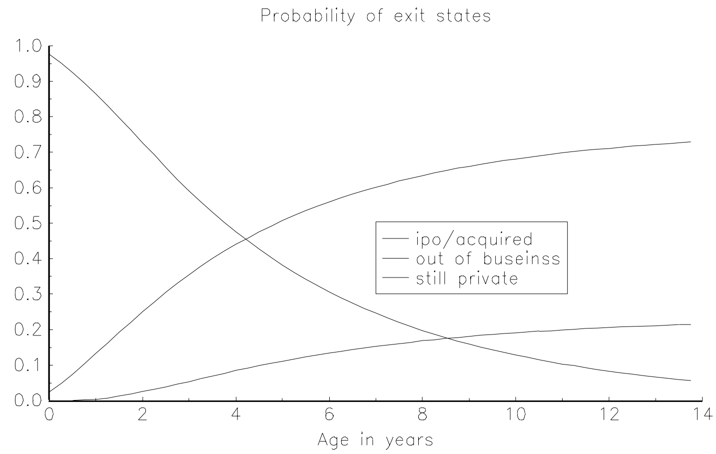

Figure 11: Cumulative probability of various exit states as a function of investment age. For example, the "ipo" line gives the chance of going public or being acquired on or before the age given on the horizontal axis.

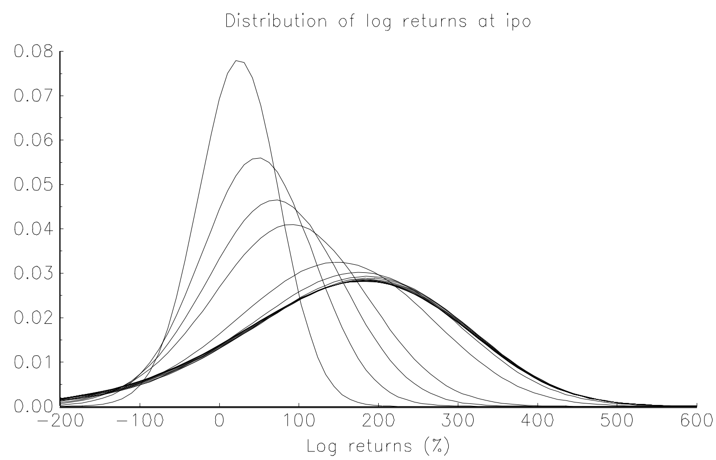

Figure 12: Model's predictions for the distribution of log returns to ipo/acquisition. Each curve represents one horizon. The highest curve gives the distibution of returns for projects that end one quarter after beginning. Marching down and to the right, the curves are at 6 months, 9 months, and then 1,2,3,...14 year horizons.

public is not a strong function of value. Thus, we essentially see the underlying value process here without a strong selection bias. However, once the value has time to build up to the levels 200%-400% that make an ipo increasingly likely in Figure 10, then the return distribution is dominated by the probability of going public. The distribution of returns to ipo approaches a *constant*, independent of horizon, as the horizon lengthens.

Comparing Figure 12 with Figure 8, you can see that the model does an excellent job of matching the return distributions, both quantitatively and qualitatively. The modal 3 year log return is about 150%, just as in Figure 8. This graph confirms that it is entirely consistent to have such a low (5.2%) average growth of log value, while typical returns to ipo are very large.

Industry breakdown

The middle set of columns of Table 5 present the same estimation broken down by industry category. Interestingly, the health (including biotech) industry generated a higher mean log return and lower standard deviation than the information technology sector. However, the higher standard deviation for IT gives it a much larger arithmetic mean. Part of the difference between health and IT comes from the larger out of business cutoff k in the health sector. ML finds it better to match the lower tail with a higher cutoff than with higher volatility and a lower cutoff. Retail has an even lower–actually negative–geometric mean, but an even larger standard deviation, giving it the largest arithmetic mean return in the group. As the *level* of the arithmetic mean comes almost entirely from the volatility rather than the geometric mean, so *differences* in arithmetic means come almost entirely from differences in volatility.

The a and b parameters governs the location of the probability of going public. a is always near one. As b increases, the curve plotted in Figure 10 flattens. Interestingly, the retail industry decisions to go public seem a much less sensitive function of value than are the other industries.

Financing round breakdown

The final set of columns in Table 5 break the estimation down by financing round. It is interesting to see if first, second, third, etc. rounds have different characteristics. In addition, this breakdown allows us to examine one source of statistical problems, the correlation of returns across rounds in the same firm.

The mean log returns decline in subsequent rounds, though this pattern is on the borderline of significance. The volatility declines as well; the lower volatility means that the average arithmetic return falls. The bankruptcy point k. This does suggest that later rounds are "less risky" than initial rounds.

The b parameter decreases dramatically as we move to later rounds.

It may be reasonable that later rounds are "more mature" projects, so investors do not stick with them through quite as much initial ill-fortune as they do in the first round. This also makes a lot of sense: later rounds require a lower increase in value before going public.

The standard error of the mean in the first round (0.99) is about 1/3 greater than the standard error of all rounds taken together (0.66), using less than half of the data points. This suggests that the standard errors calculated ignoring correlation between multiple rounds in the same firm were not orders of magnitude too optimistic.

5.2.2 Alphas and betas

Table 7 presents maximum likelihood estimates of the market model in logs,

$$\ln\left(\frac{V_{t+\Delta}}{V_t}\right) = \gamma + \ln R_t^f + \delta(\ln R_{t+\Delta}^m - \ln R_t^f) + \varepsilon_{t+\Delta}.$$

I used three reference portfolios, the S&P500, the Nasdaq and the smallest Nasdaq decile. While the latter models are not a CAPM, we can use them to compare the performance of private equity to a portfolio of comparable publicly traded stocks. Table 8 presents the less interesting estimates of k (bankruptcy cutoff), a, b (probability of going public as a function of value). The preferred estimates are marked "End of sample" in the "Exit date" column. I focus on those first, and explain the difference later.

Relative to the S&P500, the venture capital investments score a modest regression intercept $\gamma = -2.7\%$ and regression coefficient $\delta = 0.80$. (Though the simulation is quarterly, I express all returns as annualized percentages.) Comparing this result with Table 3, we see that correcting for selection bias makes an important difference. In Table 3, the regression in logs showed an intercept γ of 92% with a regression coefficient δ of 0.37. Since the unconditional volatility was so large, it is no surprise that the residual is large as well, with a 97% standard deviation.

The next three columns of Table 7 give the implied intercept and regression coefficient for a CAPM in *levels*. In the continuous time limit, $\beta = \delta$, but the α intercept is different from γ, so I report that calculation as α_c. If we use a discrete time lognormal model, both α and β are different from γ and δ, and I report that calculation as α_d and β_d. (I derive the formulas in the Appendix.) The major difference between logs and levels is that the intercepts in levels, α, add $1/2\sigma^2$ to the intercepts in logs, γ, just as arithmetic average returns add $1/2\sigma^2$ to the geometric average. As a result, even though the log intercepts are negative, when we add nearly 1/2 of a nearly 100% standard deviation to them, we obtain astonishingly large intercepts in levels, around 40%.

The results using the total NASDAQ return and the NASDAQ smallest decile are surprisingly similar to the results using the S&P500, given that these indices are rather poorly correlated. The NASDAQ mean return is slightly larger, so, with about the same beta, the intercepts are slightly smaller.

Index	Exit date	γ	(s.e.)	δ	(s.e.)	σ	(s.e.)	α_c	α_d	β_d
S&P500	End of sample	-2.7	(0.8)	0.80	(0.002)	97.3	(2.5)	44.4	47.6	0.88
Nasdaq	End of sample	-7.1	(1.2)	0.90	(0.02)	95.2	(2.5)	38.0	40.4	0.98
Nasdaq small	End of sample	-4.3	(1.0)	0.71	(0.02)	96.0	(2.5)	40.8	43.6	0.77
S&P500	On or before	5.3	(1.3)	0.92	(0.04)	91.6	(2.1)	47.5	51.1	1.03
Nasdaq	On or before	5.9	(0.9)	1.16	(0.02)	80.4	(3.0)	38.8	41.1	1.29
Nasdaq small	On or before	3.5	(0.3)	0.67	(0.02)	90.9	(2.4)	43.7	46.9	0.73
S&P500	As is	6.5	(1.2)	-0.38	(0.04)	101.2	(2.4)	58.3	63.5	-0.43

Table 7. Maximum likelihood estimates of the model

$$\ln\left(\frac{V_{t+\Delta}}{V_t}\right) = \ln R_t^f + \gamma + \delta(\ln R_{t+\Delta}^m - \ln R_t^f) + \varepsilon_{t+\Delta}.$$

γ, σ and α are presented in annual percentage return units. The parameters of the implied CAPM in levels α_c, α_d, β_d are calculated by (11) (13) (12), presented in the appendix. "End of sample" estimates ignore out-of-business dates; they calculate the likelihood from the probability that a firm has gone out of business at some point before the end of the sample. "On or before" estimates treat the out of business date as an upper bound; they calculate the likelihood from the probability that the firm has gone out of business at some point on or before the reported out of business date. "As is" estimates treat the out of business date as real; they use the probability that the firm has gone out of business in the reported quarter.

Index	Exit date	$k(\%)$	s.e.	a	s.e.	b	s.e.
S&P500	End of sample	9.8	(0.6)	0.78	(0.03)	4.79	(0.14)
Nasdaq	End of sample	10.9	(0.6)	0.76	(0.04)	4.93	(0.16)
Nasdaq small	End of sample	9.8	(0.6)	0.82	(0.03)	4.59	(0.11)
S&P500	On or before	18.1	(0.6)	0.50	(0.09)	7.25	(0.61)
Nasdaq	On or before	30.0	(0.8)	0.25	(0.06)	14.0	(0.71)
Nasdaq small	On or before	16.5	(0.6)	0.65	(0.02)	5.70	(0.11)
S&P500	As is	6.0	(0.3)	0.87	(0.02)	4.33	(0.89)

Table 8. Maximum likelihood estimates of the remaining parameters for the model described in Table 7. k is the cutoff value for going out of business, and a and b describe the probability of going public as a function of value. For all runs, the parameters describing the probability of observing good data given out of business and ipo/acquisition are $c = 0.95, d = 0.51$

Table 9 presents a breakdown of results by financing round. As usual, the log intercepts are small. Interestingly, the regression slopes δ and β decline uniformly from near one to near zero as we progress to later financing rounds. Later financing rounds are much less sensitive to market conditions. The residual volatility also decreases as we move to later financing rounds. The alphas decline as well, though some of the decrease in expected arithmetic return for the later financing rounds is countered by a decreased beta, so the alphas do not decline as much as the expected returns did. The k cutoffs for abandoning a project increase as they did before; investors give up earlier on later rounds. Again, this finding is linked to the finding of lower volatility.

Round	γ	δ	σ	α_c	α_d	β_d	k	a	b	c	d
1	0.5	0.89	99.0	49.7	53.7	1.004	9.0	0.98	4.47	0.93	0.41
(s.e.)	(1.5)	(0.05)	(3.0)				(0.6)	(0.02)	(0.07)		
2	-4.2	0.70	100.4	46.0	49.4	0.78	8.2	1.06	6.62	0.97	0.54
(s.e.)	(1.3)	(0.02)	(3.5)				(0.6)	(0.45)	(0.13)		
3	-0.74	0.38	82.1	32.7	34.6	0.41	16.5	1.01	3.35	0.98	0.63
(s.e.)	(2.8)	(0.07)	(4.28)				(1.0)	(0.07)	(0.21)		
4	0.98	0.17	86.8	38.5	41.0	0.18	16.5	0.85	3.57	0.96	0.68
(s.e.)	(3.5)	(0.12)	(3.5)				(0.3)	(0.11)	(0.36)		

Table 9. Estimates of the log market model broken down by financing rounds. The specification is

$$\ln\left(\frac{V_{t+\Delta}}{V_t}\right) = \ln R_t^f + \gamma + \delta(\ln R_{t+\Delta}^m - \ln R_t^f) + \varepsilon_{t+\Delta}.$$

Each case uses the S&P500 return for R^m and three month T bill rate for R^f. All estimates ignore out of business dates, i.e. calculate the probability of going out of business on or before the end of the sample.

Stylized facts behind the estimates, and handling the out-of-business dates

The last row of Table 7 and 8 describes the most natural estimate. When a firm goes out of business, I calculate the likelihood in these rows from the probability that the firm goes out of business on the reported date. However, this estimate produces a *negative* beta of -0.38! Before accepting this estimate, we need to see what stylized fact drives it, as well as the more successful estimates.

Figure 13 presents the percentage of outstanding rounds that go public each quarter, together with the previous year's return on the S&P500. You can see the clear pattern – firms go public in up markets. This pattern should produce a positive beta estimate–a rising market and a positive beta pushes more firms over the edge to ipo. This fact is behind the positive beta estimate in Table 3, that used only the ipos. Figure 14 presents the corresponding graph for firms that are acquired. Here, the pattern is much weaker. Still, the estimate combines acquisitions and ipos, so we do not see an explanation for a negative beta estimate. Figure 15 presents the fraction that go out of business. Here we see a surprising pattern. The data record two huge waves of firms going out of business. Furthermore, these waves come on the heels of *positive* market returns. In the model, firms go out of business when their value *declines* below k. This is the stylized fact behind the negative beta estimate.

Figure 16 digs a little deeper and shows the fraction of rounds that go out of business on or before each date. As the figure shows, a large fraction of rounds go out of business in two weeks, one in February 1995 and one in September 1997. This looks suspiciously like a data error–did 25% of all venture capital investments made between 1987 and June 2000 cease operations in a single week in September 1997? However, conversations with VentureOne have not helped us to track down the story behind these surprising dates.

We could treat the data on out-of-business dates as an upper bound – the firm went out of business on or before the indicated date. Perhaps VentureOne caught up with a stock of out of business rounds in two big waves. The "On or before" estimates of Table 7 and 8 treat the dates in this way. The resulting intercepts are somewhat higher, changing from about negative 5% to about positive 5%. The volatilities are still large so the implied intercepts in levels are still huge.

However, we did learn from conversations with VentureOne that when there is no other date information, they report the out of business date as the last date at which the firm was known to be *in* business. This suggests that the date is not an upper bound, so that all we really know is that the firm went out of business at some point in the sample. The estimates marked "End of sample" treat the out of business dates this way.

It is unfortunate to ignore so much sample information about *when* a firm goes out of business. In particular, when studies such as this one are extended to the recent period in which the NASDAQ fell dramatically and a wave of firms going out of business followed, this information will refine and possibly change substantially our estimates of both slopes and intercepts. However, it is clear that a researcher will have to devote a lot of effort to measuring the dates at which firms go out of business in order to use that information.

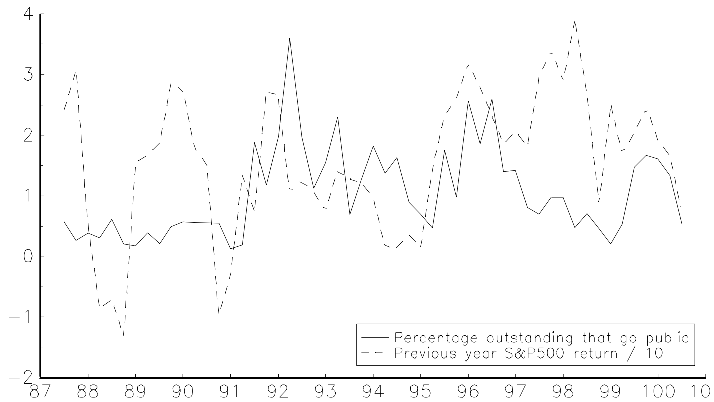

Figure 13: Percentage of outstanding rounds that go public each quarter, and the previous year's return on the S&P500.

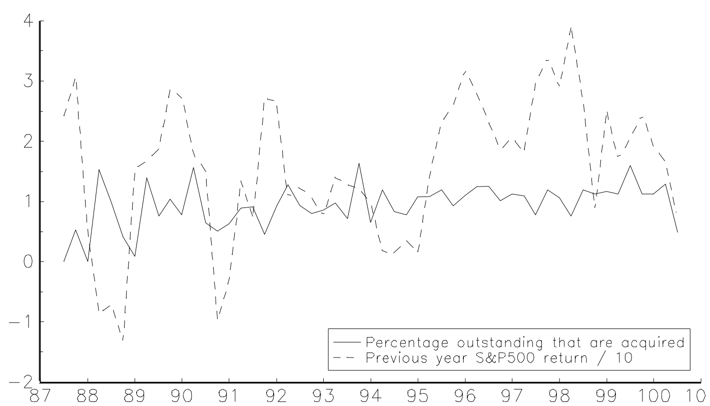

Figure 14: Percentage of outstanding rounds that are acquired each quarter together with the previous year's return on the S&P500

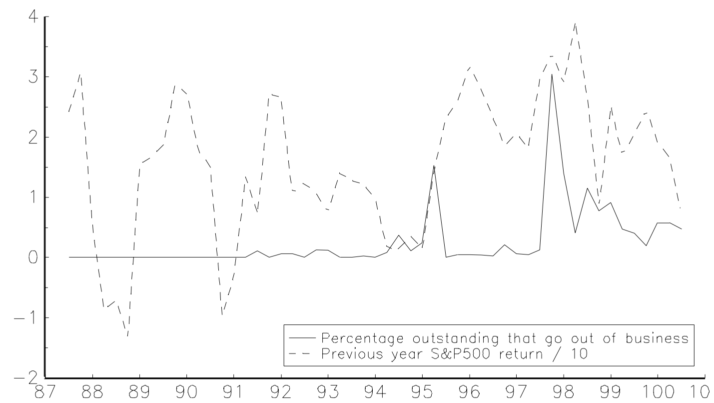

Figure 15: Fraction of outstanding rounds that go out of business each quarter, together with the previous year's S&P500 return.

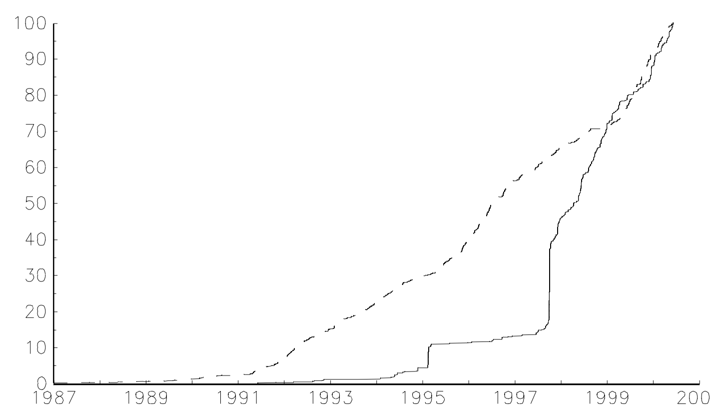

Figure 16: Percentage of rounds that have gone out of business at each date (solid line) and percentage of rounds that have gone public at each date (dashed line).

6 Implications

The mean, variance and intercept of log returns are sensible, but the volatility gives rise to large arithmetic average returns and alphas. I consider here what this means, how it could have come out differently, and what it does or does not imply.

Inescapable means, volatility and alphas

Figure 17 shows the distribution of a lognormal with mean log return $\mu = 0$ and $\sigma = 100\%$. The mean arithmetic return is $100 \times (e^{1/2} - 1) = 64\%$. As you can see though, that mean comes from a very large probability of losing money, and a much smaller probability of a dramatic gain. VC investments are very much like options. Volatility is good – it raises the chance of the large payoff, without greatly increasing the chance of a poor return. You can't do worse than lose your initial investment.

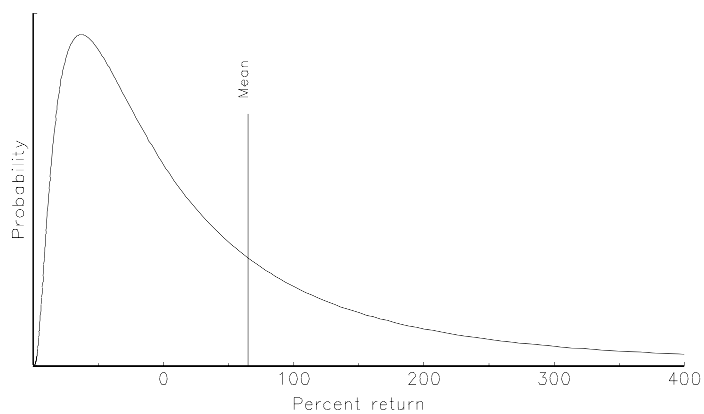

Figure 17: Distribution of a lognormal return with mean log return $\mu = 0$ and standard deviation of log return $\sigma = 100\%$. The vertical line shows the mean arithmetic return.

With this figure in mind, we can think about how the estimates could have come out differently. One possibility is that we could have estimated a 40 to 50% negative expected log return. This change would squash the peak of Figure 17 even more to the left. VC investments do not lose money quite so regularly. Such a process would lead to far more frequent failures, which is why the ML estimate settles on a higher value. To check, I started the estimation off with a 50% negative mean log return, and it came back to the same estimates.

I tried specifying a normal rather than a lognormal distribution, in which case the arithmetic and geometric averages are the same. This specification is totally at odds with the data. There are a few data points with spectacular returns in a short time period. For example, one round resulted in a factor of 290 increase in value (29,000%) in 3 months. Such a rare event is possible with a lognormal – $\ln(290) = 5.6$, which is 11 standard deviations above zero with a quarterly 0.5 standard deviation. But such an event is far beyond the range that we can even calculate a probability with a normal distribution, unless we raise the volatility to thousands of percent. But if we raise the volatility that much, then half of outstanding projects must go under each period, and the data do not show this. A lognormal, or even a fatter-tailed distribution, is necessary to capture occasional dramatic positive returns, and the limited number of failures in the data.

If volatility had been lower, the arithmetic averages would have come out much more like the sensible geometric averages. The volatility is identified by the speed with which projects go either bankrupt or to ipo, as well as by the magnitude of returns when projects do go public.. To some extent, ML trades volatility for estimates of the lower bound k. For example, at my initial guess for the bankruptcy point $k = 30\%$, the best estimate of volatility is about 60% rather than 100%. The higher bankruptcy point then restores the probability of bankruptcy. However, the overall likelihood, matching the exit probabilities at all different horizons, is larger if one uses larger volatility and a much lower cutoff $k = 5.4\%$.

The high volatility estimate is also nearly inevitable given the high mean and volatility of returns when there is an ipo. The ipo end of the selection bias reduces the volatility of observed returns compared to true returns. If every firm goes public at a return of 1,000%, then the standard deviation of observed returns is zero, no matter what the standard deviation of actual returns. In more modest cases such as Figure 1, the true return distribution gets squashed by the probability of going public. Thus, the 138% standard deviation of log returns to ipo we saw in the data will almost surely generate a similar standard deviation in the model. Given this logic, it is if anything surprising that we recover a volatility estimate *lower* than 138%. The estimated probability of going public is a surprisingly slowly rising function of value. If we estimate or impose a more sharply rising function (for example, by separating ipos and acquisitions), this will lower measured volatility given underlying volatility, and thus require an even higher estimate of underlying volatility.

In sum, in order to fit occasional spectacular returns, the limited fraction that go out of business, and the large volatility of returns to ipo, the large volatility is inescapable, the skewed distribution is inescapable, and a dramatically lower mean log return is quite unlikely. The only other possibility is that beta is really something like 3 rather than the 1 that I have estimated. Here, having to throw out the date at which firms go out of business is particularly unfortunate. If the true, correctly measured pattern, shows many failures after market declines, the beta estimate will be substantially raised. Future studies that have better data and can include the NASDAQ crash and dot-com shakeout may settle on higher beta estimates.

Portfolios

The portfolio implications of large expected returns and alphas are not so obvious as they seem initially, because the volatility is so huge.

1. An individual VC investment is not particularly attractive, despite the high

average returns and alphas. In my sample, the Sharpe ratio of a single VC investment is approximately half of the S&P500 Sharpe ratio in the same period. Furthermore, a single VC investment is far from normally distributed, as dramatized by Figure 17. Sharpe ratios are a bad way to evaluate such investments. A log utility investor ranks portfolios by $E(\ln R)$ directly. The average log return of a single VC investment is about half that of the S&P500. (Table 5)

2. Adding a single VC investment to a market portfolio does not give a huge increase in performance, because the residual volatility of VC investments is so large. To make the point, Figure 18 calculates the in-sample Sharpe ratio of an investment in the S&P500 and one VC round. The maximum Sharpe ratio occurs with only 4% of wealth placed in the VC firm, and is barely higher than the Sharpe ratio of the S&P500 alone. The reason, of course, is the tremendous volatility of the VC investment. Even though the alpha is positive, so that an optimal portfolio puts *some* weight on the VC investment, as soon as you put any substantial weight on that investment, portfolio volatility rises dramatically[3].

3. Of course the promise of alpha is that a *well-diversified* portfolio of many high α investments should yield spectacular results. But now we are on thinner ground. If the residuals are independent *of each other* – if $E(\varepsilon_i \varepsilon_j) = 0$ – then one can achieve an arbitrarily high Sharpe ratio with a sufficient number of small VC investments. But we do not know this. I did not estimate any correlation structure between VC investments. It's quite possible that there is a strong common component to VC investments, so that a "well-diversified" portfolio is still quite volatile. In the Fall of 2000, many VC investments went out of business at the same time, and many more were substantially delayed, all at the same time. The smell of a common component is there. Furthermore, VC investments have until very recently been quite difficult to diversify since they were structured as limited partnerships.

4. A venture capital investment is illiquid. If the market goes down, not only will

[3] I calculated Figure 18 as

$$\text{Sharpe ratio} = \frac{E(R^p) - R^f}{\sigma(R^p)}$$

$$\begin{aligned} E(R^p) - R^f &= w\left[E(R) - R^f\right] + (1-w)\left[E(R^{S\&P}) - R^f\right] \\ &= wE\left[\alpha + \beta(R^{S\&P} - R^f)\right] + (1-w)\left[E(R^{S\&P}) - R^f\right] \\ & \quad w\alpha + \left[w\beta + (1-w)\right]\left[E\left(R^{S\&P}\right) - R^f\right] \end{aligned}$$

$$\begin{aligned} \sigma(R^p) &= \sigma\left[wR + (1-w)R^{S\&P}\right] \\ &= \sigma\left[w\left(\alpha + R^f + \beta(R^{S\&P} - R^f) + \varepsilon\right) + (1-w)R^{S\&P}\right] \\ &= \sigma\left[\left(1 + w\left(\beta - 1\right)\right)R^{S\&P} + w\varepsilon\right] \\ &= \sqrt{\left(1 + w\left(\beta - 1\right)\right)^2 \sigma^2\left(R^{S\&P}\right) + w^2\sigma^2\left(\varepsilon\right)}. \end{aligned}$$

I used the numerical values for α and β from Table 7, the values for the S&P return from Table 5 and a 5% risk-free rate.

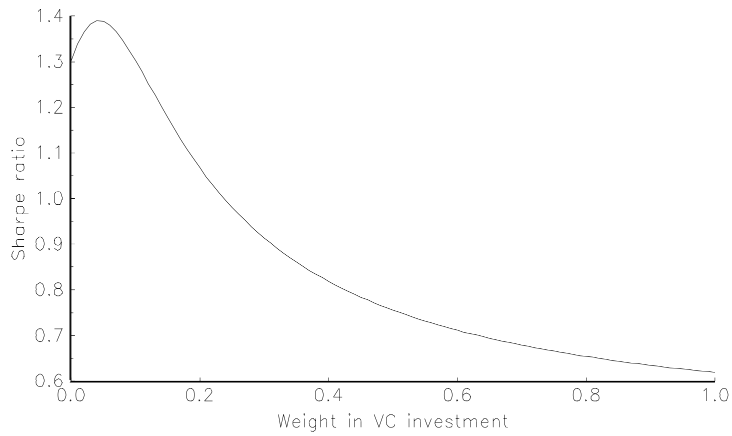

Figure 18: Sharpe ratio attained by an investment in the S&P500 and one venture capital round.

returns be lower, but they may be more delayed. Standard portfolio theory with a fixed horizon and/or constantly tradeable assets does not necessarily apply, even if the alphas are large.

7 Conclusions and extensions

In sum, the selection bias correction neatly accounts for the *log* returns. It reduces the mean log return from 100% or more to a sensible 5%; it reduces the intercept in a log market model from 93% to near zero. However, the huge volatility of log returns and the market model regression means that arithmetic returns and alphas are still very large, in the range of 40 to 50%.

There are many ways that this work can be extended, though each involves a substantial investment in programming and computer time. My model of the ipo and acquisition process is very stylized–I assumed that ipo, acquisition, and going out of business were only a function of the firm's value at a point in time. Most easily, one might separate ipo and acquisition, at the (not insubstantial) cost of two more parameters. More ambitiously, the decision to go public may well depend on the market as well as on the value of the particular firm. There do seem to be waves of ipos in "good markets," high prices relative to dividends, book values or earnings. While such waves will also raise the value of a particular firm, it may be the case that firms are more likely to go public, even given their own values, in high stock markets. Finally, age and industry effects are likely in all of these decisions.

Multiple risk factors are an obvious generalization, though with this approach each additional regressor multiplies the simulation time dramatically. Combining the two modifications, the risks (betas, standard deviation) of the firm are also likely to change as its value increases, as the breakout by financing round suggests.

The modeling philosophy can be extended to consider multiple rounds in the same firm more explicitly. The probability of needing additional rounds will depend on value and other parameters, though, so this modification will also introduce substantial complexity and extra parameters.

Most importantly, we will only get better results with better data. Establishing the dates at which firms go out of business is important to this estimation procedure.

8 References

Bygrave, William D. and Jeffrey A. Timmons, 1992, *Venture Capital at the Crossroads* Boston: Harvard Business School Press.

Gompers, Paul A., and Josh Lerner, 1997, "Risk and Reward in Private Equity Investments: The Challenge of Performance Assessment," Journal of Private Equity (Winter 1997): 5-12.

Gompers, Paul A., and Josh Lerner, 2000, "Money Chasing Deals? The Impact of Fund Inflows on Private Equity Valuations," *Journal of Financial Economics* 55, 281-325.

Long, Autsin M. III, 1999, "Inferring Period Variability of Private Market Returns as Measured by σ from the Range of Value (Wealth) Outcomes over Time, *Journal of Private Equity* 5, 63-96.

Moskowitz, Tobias J. and Annette Vissing-Jorgenson, 2000, "The Private Equity Premium Puzzle," Manuscript, University of Chicago.

Reyes, Jesse E., 1997, "Industry Struggling to Forge Tools for Measuring Risk," *Venture Capital Journal*Venture Economics, *Investment Benchmarks: Venture Capital*

Smith, Janet Kiholm and Richard L., 2000, *Entrepreneurial Finance* New York: Wiley and Sons

Venture Economics, 2000, Press release, May 1, 2000 at www.ventureeconomics.com

9 Appendix

9.1 Details of data selection

Details of the construction of Figures 2 and 3

I removed rounds that ended in ipo, acquisition or bankruptcy if they had missing dates. I also removed cases in which the end date was earlier than the start date, cases in which the fate was unknown, and start dates 19870101, which codes for all values earlier than this date.

To estimate the fractions at, say, a 4 year age, I started with all rounds with a start date earlier than 4 years before the end of the sample – rounds that had a chance to achieve the 4 year age before the end of the sample. The fraction out of business is then the fraction of all these rounds that went out of business at an age less than or equal to 4 years. For example, a round that started in Jan 1991 and went out of business in June 1993 would be counted.

There is a selection bias with this measure: only firms that go out of business, are acquired or go public can have bad exit dates; firms that are still private cannot be removed from the sample for bad exit dates, and so are overrepresented. To account for this bias, I calculated the out of business fraction at, say, 4 years, as

$$\frac{\text{out}_4 \times \text{out ratio}}{\text{total}_4 \times \text{total ratio}}$$

where

> $\text{out}_4 =$ number of rounds that started more than 4 years before the end of the sample, and went out of business in less than or equal to 4 years
>
> $\text{total}_4 =$ number of rounds that started more than 4 years before the end of the sample.
>
> out ratio = number of out-of-business rounds after selections/ number of out-of-business rounds before selections
>
> total ratio = total number of rounds after selection / total number of rounds before selection .

I followed the same procedure to reweight the ipo or acquired category. I calculated the "still private" category as one less the last two categories. The weights do not make a difference noticeable to the eye in Figures 2 and 3.

Details of return selection

To compute Tables 2-4 and Figures 2-9, I selected the data as follows. I removed all rounds in which the value information implied that the VC investors owned more than 100% or less than or equal to 0% of the company. (For example, if $10 million is raised, and the post-round valuation is \$5 million). I removed observations with ipo or acquisition less than two months after the financing round date. I eliminated observations that had missing financing round dates, missing ipo or acquisition dates, or missing return data. I also removed observations with round date 19870101, which codes for unknown financing date before this date.

9.2 Logs to levels in the CAPM

This section derives the formulas for $\alpha_c, \alpha_d,\ \beta_d$ in Table 7. From the estimated market model in logs,

$$\ln\left(\frac{V^i_{t+\Delta}}{V^i_t}\right) - \ln R^f_t = \gamma + \delta(\ln R^m_{t+\Delta} - \ln R^f_t) + \varepsilon^i_{t+\Delta}. \tag{10}$$

We want to find the implied CAPM in levels, i.e.

$$\frac{V^i_{t+\Delta}}{V^i_t} - R^f_t = \alpha + \beta(R^m_{t+\Delta} - R^f_t) + v^i_{t+\Delta}.$$

Results. In the continuous time limit, $\beta = \delta$ and $\sigma(\varepsilon) = \sigma(v)$, but

$$\alpha_c = \gamma + \frac{1}{2}\delta\left(\delta - 1\right)\sigma^2_m + \frac{1}{2}\sigma^2. \tag{11}$$

As advertized, the major effect is a familiar $1/2\sigma^2$ term. If we model the returns as lognormals in discrete time, we obtain instead

$$\begin{aligned}
\beta_d &= e^{\gamma+(\delta-1)\left(E(\ln R^m)-\ln R^f\right)+\frac{1}{2}\sigma^2+\frac{1}{2}\left(\delta^2-1\right)\sigma^2_m}\frac{\left(e^{\delta\sigma^2_m}-1\right)}{\left(e^{\sigma^2_m}-1\right)} && (12)\\
\alpha_d &= e^{\ln(R^f)}\left\{\left(e^{\gamma+\delta\left(E(\ln R^m)-\ln R^f\right)+\frac{1}{2}\delta^2\sigma^2_m+\frac{1}{2}\sigma^2}-1\right)-\beta\left(e^{\left(\mu_m-\ln R^f\right)+\frac{1}{2}\sigma^2_m}-1\right)\right\} && (13)
\end{aligned}$$

Algebra for continuous-time limit. We start with the continuous time version of the log market model,

$$\begin{aligned}
d\ln V &= \left(r^f+\gamma\right)dt + \delta\left(d\ln P^m - r^f dt\right) + \sigma dz\\
d\ln P^m &= \mu_m dt + \sigma_m dz^m\\
E(dz dz^m) &= 0
\end{aligned}$$

Substituting,

$$\begin{aligned}
d\ln V &= \left(r^f(1-\delta)+\gamma\right)dt + \delta\left(\mu_m dt + \sigma_m dz^m\right) + \sigma dz\\
&= \left(r^f+\gamma+\delta\left(\mu_m - r^f\right)\right)dt + \delta\sigma_m dz^m + \sigma dz
\end{aligned}$$

Now, we can transform to levels. Using Ito's lemma,

$$\begin{aligned}\frac{dV}{V} &= \frac{d\left(e^{\ln V}\right)}{V} = d\ln V + \frac{1}{2}d\ln V^2 \\ &= \left[r^f + \gamma + \delta\left(\mu_m - r^f\right)\right]dt + \delta\sigma_m dz^m + \sigma dz + \frac{1}{2}\left(\delta^2\sigma_m^2 + \sigma^2\right)dt\end{aligned}$$

$$\frac{dP^m}{P^m} = \left(\mu_m + \frac{1}{2}\sigma_m^2\right)dt + \sigma_m dz^m$$

Using the latter expression to substitute for dz^m,

$$\frac{dV}{V} = \left[r^f + \gamma + \delta\left(\mu_m - r^f\right) + \frac{1}{2}\left(\delta^2\sigma_m^2 + \sigma^2\right)\right]dt + \delta\frac{dP^m}{P^m} - \delta\left(\mu_m + \frac{1}{2}\sigma_m^2\right)dt + \sigma dz$$

or, finally,

$$\frac{dV}{V} - r^f dt = \left[\gamma + \frac{1}{2}\delta\left(\delta - 1\right)\sigma_m^2 + \frac{1}{2}\sigma^2\right]dt + \delta\left(\frac{dP^m}{P^m} - r^f dt\right) + \sigma dz$$

We see that $\beta = \delta$, and the errors are the same, but we derive formula (11) relating the log intercept to the intercept in levels.

Algebra for the discrete-time lognormal calculation. From the model (10), we want to find the implied regression in levels,

$$R_{t+\Delta} - R_t^f = \alpha + \beta(R_{t+\Delta}^m - R_t^f) + \varepsilon_{t+\Delta}^i$$

where

$$R_{t+\Delta} \equiv \frac{V_{t+\Delta}}{V_t}.$$

(It does not matter that the conditional expectation of $V_{t+\Delta}/V_t$ is a nonlinear function of R^m. The CAPM specifies the projection or linear regression.) We start with beta,

$$\beta = \frac{cov\left[R_{t+\Delta}, R^m\right]}{var(R^m)}.$$

The denominator is

$$\begin{aligned} var(R^m) &= E\left(R^{m2}\right) - \left[E(R^m)\right]^2 \\ &= E\left[e^{2\ln R^m}\right] - \left[E\left(e^{\ln R^m}\right)\right]^2 \\ &= e^{2\mu_m + 2\sigma_m^2} - \left(e^{\mu_m + \frac{1}{2}\sigma_m^2}\right)^2 \\ &= e^{2\mu_m + 2\sigma_m^2} - e^{2\mu_m + \sigma_m^2} \\ &= e^{2\mu_m + \sigma_m^2}\left(e^{\sigma_m^2} - 1\right). \end{aligned}$$

The numerator is

$$cov\left(R,\ R^m\right) = E\left[RR^m\right] - E(R)E(R^m) = E\left[E(R|R^m)R^m\right] - E\left[E(R|R^m)\right]E(R^m)$$

Now,

$$E(R|R^m) = E\left\{e^{\gamma+\ln R^f+\delta\left[\ln R^m-\ln R^f\right]+\varepsilon^i}|R^m\right\} = e^{\gamma+\ln R_t^f+\delta\left[\ln R^m-\ln R^f\right]+\frac{1}{2}\sigma^2}.$$

Thus,

$$\begin{aligned}
cov\left(R,\ R^m\right) &= E\left[e^{\gamma+\ln R^f+\delta\left(\ln R^m-\ln R^f\right)+\frac{1}{2}\sigma^2+\ln R^m}\right] - E\left[e^{\gamma+\ln R^f+\delta\left(\ln R^m-\ln R^f\right)+\frac{1}{2}\sigma^2}\right] E\left[e^{\ln R^m}\right] \\
&= E\left[e^{\gamma+(1-\delta)\ln R^f+(1+\delta)\ln R^m+\frac{1}{2}\sigma^2}\right] - E\left[e^{\gamma+(1-\delta)\ln R^f+\delta\ln R^m+\frac{1}{2}\sigma^2}\right] E\left[e^{\ln R^m}\right] \\
&= e^{\gamma+(1-\delta)\ln R^f+(1+\delta)\mu_m+\frac{1}{2}(1+\delta)^2\sigma_m^2+\frac{1}{2}\sigma^2} - e^{\gamma+(1-\delta)\ln R^f+(1+\delta)\mu_m+\frac{1}{2}\left(\delta^2+1\right)\sigma_m^2+\frac{1}{2}\sigma^2} \\
&= e^{\gamma+(1-\delta)\ln R^f+(1+\delta)\mu_m+\frac{1}{2}\sigma^2}\left(e^{\frac{1}{2}(1+\delta)^2\sigma_m^2} - e^{\frac{1}{2}\left(\delta^2+1\right)\sigma_m^2}\right) \\
&= e^{\gamma+(1-\delta)\ln R^f+(1+\delta)\mu_m+\frac{1}{2}\sigma^2}\left(e^{\frac{1}{2}(1+2\delta+\delta^2)\sigma_m^2} - e^{\frac{1}{2}\left(1+\delta^2\right)\sigma_m^2}\right) \\
&= e^{\gamma+(1-\delta)\ln R^f+(1+\delta)\mu_m+\frac{1}{2}\sigma^2+\frac{1}{2}(1+\delta^2)\sigma_m^2}\left(e^{\delta\sigma_m^2}-1\right).
\end{aligned}$$

Putting it all together,

$$\begin{aligned}
\beta &= \frac{e^{\gamma+(1-\delta)\ln R^f+(1+\delta)\mu_m+\frac{1}{2}\sigma^2+\frac{1}{2}(1+\delta^2)\sigma_m^2}\left(e^{\delta\sigma_m^2}-1\right)}{e^{2\mu_m+\sigma_m^2}\left(e^{\sigma_m^2}-1\right)} \\
&= e^{\gamma+(\delta-1)\left(\mu_m-\ln R^f\right)+\frac{1}{2}\sigma^2+\frac{1}{2}\left(\delta^2-1\right)\sigma_m^2}\frac{\left(e^{\delta\sigma_m^2}-1\right)}{\left(e^{\sigma_m^2}-1\right)}.
\end{aligned}$$

Continuing for α,

$$\begin{aligned}
\alpha &= E(R) - R^f - \beta E(R^m) - R^f \\
&= E\left[E\left(R|R^m\right)\right] - R^f - \beta\left[E(R^m) - R^f\right] \\
&= e^{\gamma+(1-\delta)\ln R^f+\delta\mu_m+\frac{1}{2}\delta^2\sigma_m^2+\frac{1}{2}\sigma^2} - e^{\ln R^f} - \beta\left[e^{\mu_m+\frac{1}{2}\sigma_m^2} - e^{\ln R^f}\right] \\
&= e^{\ln R^f}\left(e^{\gamma+\delta\left(\mu_m-\ln R^f\right)+\frac{1}{2}\delta^2\sigma_m^2+\frac{1}{2}\sigma^2}-1\right) - \beta\left(e^{\mu_m+\frac{1}{2}\sigma_m^2} - e^{\ln R_t^f}\right) \\
&= e^{\ln R^f}\left\{\left(e^{\gamma+\delta\left(\mu_m-\ln R^f\right)+\frac{1}{2}\delta^2\sigma_m^2+\frac{1}{2}\sigma^2}-1\right) - \beta\left(e^{\left(\mu_m-\ln R^f\right)+\frac{1}{2}\sigma_m^2}-1\right)\right\}
\end{aligned}$$