Dow Jones Venture Capital Index Handbook

Dow Jones Indexes (a/k/a CME Group Index Services LLC)

In Consultation with Sand Hill Econometrics

Dow Jones Indexes
A CME Group Company

TABLE OF CONTENTS

1. INTRODUCTION AND BACKGROUND

This document is published by Dow Jones Indexes, the marketing name and a licensed trademark of CME Group Index Services LLC. This document describes the calculation methodology of the Dow Jones Venture Capital Index (the "Index"). Dow Jones Indexes publishes the Dow Jones Venture Capital Index pursuant to a License and Joint Marketing Agreement with Sand Hill Econometrics, Inc. ("Sand Hill").

The Index seeks to measure the monthly value of a value-weighted portfolio of venture-capital-funded companies. As further described below, the monthly value is based on a combination of values reported by the individual venture-capital-funded companies to the Dow Jones VentureSource ("DJVS") database and/or estimates using a methodology described herein. The DJVS is owned and published by VentureOne Corporation, a subsidiary of Dow Jones & Company, Inc., and has been licensed for use in connection with the Index.

This approach is an alternative to indices that report averages of returns on investments in venture capital ("VC") funds themselves. The fund returns approach suffers from biases resulting from the difficulty of obtaining results from venture capital funds. There is no general public reporting of fund results, and the results most difficult to obtain are typically the worst results. Omission of the worst results from the averages leads to an index that is biased upward.

Unlike indices built from fund returns (the components of which may invest in a mixture of buyouts, other types of alternative assets, and public equities, as well as venture capital deals), our approach focuses on individual companies and measures the change in value (before fees and carry) for them. The Index seeks to measure venture-funded company value. The companies included in the Index are a subset of the companies in the DJVS database. We include only venture-backed companies headquartered in the United States. In addition, because companies are more easily classified into industry, stage, and other groupings than are funds, cross-sectional analysis and specialized indices are feasible.

Inclusion in the DJVS database is in part qualified by the following DJVS criteria:

"VentureSource tracks privately held and innovative companies that receive cash-for-equity financing directly from an institutional venture capital limited partnership or LLC or from another private equity entity making an equity investment structured like a venture capital round."

Venture-funded firms are privately owned, are not registered (with the SEC), and do not have shares traded in any public market. Nonetheless, there are events that give rise to market prices for the firms (i.e., a value based on an arm's-length transaction) from time to time. Venture-backed firms receive their funding intermittently over the course of several years in distinct funding "rounds," at which time a proportion of the firm's ownership is transferred to the investing VC funds in exchange for cash. The proportion of the firm's equity transferred in exchange for the cash investment is based on a negotiated "pre-money" value. A market value for the firm is thereby established at the time of each financing round.

Similarly, a value is established at the time a company goes public, is acquired, or shuts down (goes out of business). These points when firm value is revealed through a market transaction (financing rounds and exit events) provide opportunities to observe firm value, and we refer to them collectively as "valuation events."

We use these observed values along with other firm characteristics to construct a monthly value-weighted index in the following steps:

1. Estimate values for valuation events where values are not revealed.
2. Interpolate firm value for months between valuation events.
3. Estimate the value of non-exited firms between the time of their last round and the index end.
4. Construct a value-weighted, continuously invested index of value for all venture-funded companies alive (from first round of funding to exit) as of each month in the index.

Each of these four steps is described in 2. Data Preparation. In addition, Appendix A provides a more detailed discussion of the index construction, with some examples of the estimation techniques applied to a specific venture-funded company.

2. DATA PREPARATION

2.1 Estimating Value for Non-Value-Revealing Events

Sources of Data

The Dow Jones VentureSource database is a collection of detailed information on the funding and exit events of venture-backed firms throughout the United States and provides our main source of firm-level data, including known values. In addition, we merge in a few supplementary data sources as described in Appendix B. In addition, Dow Jones Indexes provides the historical public market index levels for Dow Jones U.S. Industry and Sub-Industry Indexes as categorized by the Industry Classification Benchmark ("ICB") codes, as well as the levels of the Dow Jones U.S. Total Stock Market Index^{SM} (TSM).

Estimation of Missing Firm Values

At the time of an IPO, the value of the company going public is always known or "revealed." Value is assumed to be zero at the time of a shutdown or failure. However, for many funding rounds and acquisitions the company value at the time of the event is not reported, shared, or revealed. For these events with missing value, we estimate firm value.

When companies raise money in private funding rounds, they nearly always indicate how much money they are raising, how much they have raised so far in their history, what industry they are in, their business stage, their location, and many other details. Only sometimes do they share the value implied by the terms of the funding. To estimate value for the companies not revealing, we use the information available about firms that have revealed values to estimate the value. We use the following variables as predictors of value.

Firm and Round Characteristic Variables

- Amount raised at the valuation event (raised)
- Amount raised in the previous history of the company (RTD)
- Firm industry: InfoTech, health, retail, other (categorical variable)
- Company business stage: startup, development, beta, clinical trials, shipping, profitable (categorical variable) indicators

- Company funding stage: seed, early, late, mezzanine (categorical variable)
- Whether company had a bridge round immediately prior to the event (categorical variable)
- Whether company has ever revealed value prior to this event (categorical variable)
- Value at most recent valuation event, if available
- Time elapsed since last valuation event

Selection Bias Correction Variables, Using "Uncensored" Data Collected by Sand Hill

- Whether the source of the value is a Sand Hill uncensored source[1]
- Interaction of censored indicator and amount raised in this round
- Interaction of censored indicator and total amount raised

Macroeconomic Variables

- Level of the Dow Jones TSM (Total Stock Market Index)
- Categorical variables indicating time (quarter) indicator for each quarter

We selected the variables to use in this estimation on the basis of availability and usefulness. Yes, there are other factors that would be useful (income, revenues, number of employees, etc.) but these are not regularly disclosed. Our research has demonstrated that the variables we do use are systematically related (in a forecasting sense) to company value.

Using only the (non-exit) funding rounds that have a pre-money value, we estimate the following regression using nonlinear least squares (NLLS) (note that some variables are interactions of several variables) using this specification:

$$\begin{aligned} V_{pre} = exp(c &+ \beta_1 LogRaised + \beta_2 LogRtd + \beta_{3,4,5} Ind_{1,2,3} + \beta_{6,7,8,9,10} Bs_{1,2,3,4,5} \\ &+ \beta_{11,12,13} Fs_{1,2,3} + \beta_{14} BrPrior + \beta_{15} Unc + \beta_{16} Unc \cdot LogRaised \\ &+ \beta_{17} Unc \cdot LogRtd + \beta_{18} NotHasLastKV + \beta_{19} LogLastKV \\ &+ \beta_{20} LogYearsLastKV + \beta_{21} LogTSM + \beta_{22,23,...} Qtr_{2-T} + \epsilon) \end{aligned}$$

where c denotes an intercept term and ϵ is the error term, and where:

$LogRaised$ is the log of the amount raised in this round

$LogRTD$ is the log of the cumulative amount of money raised excluding the current round

$Ind_{1,2,3}$ are the industry indicators: health, retail, other

$Bs_{1,2,3,4,5}$ are the business stage indicators: Startup, Development, Beta, Clinical Trials, Profitable

$Fs_{1,2,3}$ are the funding stage indicators: Early, Late, Mezz

$BrPrior$ is the indicator for whether the immediately prior event was a bridge round

[1] An company value is considered "uncensored" if the value is neither provided to DJVS nor generally made public, but would be available to an investor in the funded company.

$NotHasLastKV$ is an indicator denoting if there is no previous funding round with a revealed (known) value

$LogLastKV$ is the log of the last revealed value prior to his event, if available

$LogYearsLastKV$ is the log of the time elapsed (in years) since the last revealed (known) value, if available

Unc is the indicator for the source of the value and is a Sand Hill uncensored source

$Unc \bullet LogRaised$ is the interaction of censored indicator and amount raised in this round

$Unc \bullet LogRTD$ is the interaction of censored indicator and total amount raised

$LogTSM$ is the log of the Dow Jones TSM (Total Stock Market Index) level

Qtr_t is the indicator for each quarter

We exclude from the regression values that are more than four standard deviations from the mean value for any five-year period (the data in the DJVS Database extends from 1990 forward through today). A more detailed description of the process used to identify and exclude outliers from the regression is presented in Appendix B.\

We believe companies that reveal value at the time of a fund-raising event are more valuable, other things being equal, than others. We address the value-revelation bias by including data on value that are in Sand Hill's possession, but are not generally publicly available, and indicating that these data are from uncensored sources. Thus, our approach to selection bias is not a correction of the types developed by McFadden and Heckman, but a direct comparison of censored with uncensored data.

Note that the coefficients corresponding to categorical variables with multiple categories (such as $Bs_{1,2,3,4,5}$) represent a set of distinct coefficients. Similarly, $QTR_{2\text{-}T}$ denotes a set of dummy (categorical) variables, one for each quarter since the beginning of the Index, omitting the first quarter. Thus $\beta_{22,23,\ldots}$ is of dimension $1\times(T-1)$. For other (non-date) categorical variables, the most numerous category is the omitted category.

Using the estimated coefficients $\widehat{\beta}$ obtained above, we calculate estimated pre-money values $\widehat{V}_{pre}$ for all funding rounds for which this value is missing. This estimate, however, is biased because of the curvature introduced by taking the logarithm of value. To correct for this bias, we estimate a "scaling factor" (SF) using the value-weighted average difference between the actual revealed pre-money value V_{pre} and the estimated pre-money value $\widehat{V}_{pre}$ as follows:

$$SF = avg(V_{pre}) / avg(\widehat{V}_{pre})$$

The final estimate of pre-money value is then estimated as $(\widehat{V}_{pre} \cdot SF)$. After all missing pre-money values (V_{pre}) have been estimated, post-money values are calculated as $V_{post} = (V_{pre} + Raised)$, the pre-money value plus the amount raised in this round.

2.2 Estimation of Firm Value for Acquisitions When Value Is Not Revealed

To estimate missing acquisition values, we perform a nonlinear least squares regression similar to the one used to estimate non-revealed round values. The specification used to estimate non-revealed acquisition values is as follows:

$$
\begin{aligned}
V_{acq} = exp(c &+ \beta_1 TSM + \beta_2 TSM^2 + \beta_3 TSM^3 + \beta_4 LogTSM \\
&+ \beta_5 LogRTD + \beta_6 LogLastKV + \beta_7 NotHasLastKV + \beta_8 ElapFirst \\
&+ \beta_9 ElapFirst^2 + \beta_{10} ElapFirst^3 + \beta_{11} ElapLast + \beta_{12} ElapLast^2 \\
&+ \beta_{13} ElapLast^3 + \beta_{14} NotHasEES + \beta_{15} EES + \beta_{16} LogEES \\
&+ \beta_{17} AcqPub + \beta_{18} BrPrior + \beta_{19} LogMonth + \beta_{20} Prof \\
&+ \beta_{21} Info + \beta_{22} LastKVGr250 + \beta_{23} HasEES \cdot AcqPub \\
&+ \beta_{24} AcqPub \cdot LogRTD + \beta_{25} HasEES \cdot AcqPub \cdot LogRTD + \epsilon)
\end{aligned}
$$

where

TSM is the level of the Dow Jones TSM Index

$LogTSM$ is the log of the level of the Dow Jones TSM index

$LogRtd$ is the log of the raised-to-date amount for the company at acquisition (in $mm)

$LogLastKV$ is the log of the last revealed value for the company prior to acquisition

$NotHasLastKV$ is a dummy variable, = 1 if there is no prior known value, else 0

$ElapFirst$ is the time in years from the company's first venture round and the acquisition date

$ElapLast$ is the time in years from the company's last venture round and the acquisition date

$NotHasEES$ is a dummy variable, = 1 if the number of employees at acquisition is unknown, else 0

EES is the number of employees at acquisition

$LogEES$ is the log of the number of employees at acquisition

$AcqPub$ is a dummy variable, = 1 if the acquiring company is a public company, else 0

$BrPrior$ is a dummy variable, denotes if the event prior to acquisition was a bridge round

$LogMonth$ is the log of the number of months since January 1990 that the acquisition occurred

$Prof$ is a dummy variable to denote if the company was in the "Profitable" business stage

$Info$ is a dummy variable to denote if the company is in the Information Technology industry

$LastKVGr250$ is a dummy variable to denote if the last known value for the company was greater than $250 million

$HasEES$ is a dummy to denote if the number of employees at acquisition is known

We once again exclude outliers from this regression, this time by dropping acquisitions for $400 million or more (we are confident that there are no acquisitions for $400 million or more for which value was not revealed). The resulting coefficients are used to calculate estimated values for acquisitions with missing values.

As in the earlier regression for rounds, we need to address the bias that results from converting from log values to level (dollar) values by estimating a "scaling factor" (SF_{acq}) using the value-weighted average difference between the actual revealed value V_{acq} and the fitted pre-money value $\widehat{V}_{acq}$ as follows:

$$SF_{acq} = avg(V_{acq}) / avg(\widehat{V}_{acq})$$

There is a further and more important bias correction: Our research indicates that firms for which the acquisition value is missing are on average worth much less than those for which a value is readily available. Our experience is that the harder a value is to find, the lower it is. To account for this phenomenon, we further adjust downward the values calculated from applying the regression coefficients to each company's individual data (money raised, time elapsed, etc.) We calibrate this adjustment factor λ (a number between zero and one) and multiply the estimated value by this adjustment factor to get the final estimated value: $(\widehat{V}_{acq} \cdot \lambda)$. For a detailed discussion of how this adjustment factor is calculated, see Appendix C.

2.3 Monthly Interpolation Between Valuation Events

Our goal is to generate a monthly pre- and post-money value for each firm in the sample between its first and last known valuation events. In months containing a valuation event, we have pre- and post-money values (i.e., either revealed or estimated). Between those months, we interpolate to estimate monthly firm values. Denote the current month by s, the month of the most recent valuation event by t, and the month of the following event (the next event after month s) by T. Also denote an industry-specific public market index—specifically, the Dow Jones Industry and Sub-Industry Index levels based on ICB (corresponding to the high-level industrial grouping for the firm in question) in month i by M_i, the post-money value by Vi, and the pre-money value by ν_i.

2.3.1 Interpolation when the firm has a value greater than zero at both events

The interpolation method used depends on whether both the post-money value at the most recent prior value (time t) and the pre-money value at the subsequent event (time T) are strictly positive. If they are, meaning neither event represents a firm failure and neither has a zero or negative estimated value (a result of estimation error), we calculate the discount factor used between the event at time t and the event at time T according to

$$\gamma = \frac{\log\left(\frac{\nu_T / M_T}{V_t / M_t}\right)}{T - t}$$

This formulation assumes a market beta of 1. To add a market beta (β) other than 1 and facilitate the use of a commonly understood β, let us first rewrite the equations above to use returns rather than levels.

Let

$$R^{v}_{t,s} = \nu_s / V_t$$

Now we rewrite the above formula for γ as

$$\gamma = \frac{\log\left(\frac{R^{\nu}_{t,T}}{R^{M}_{t,T}}\right)}{T - t}$$

and

$$R^{\nu}_{t,s} = R^{M}_{t,s} e^{\gamma(s-t)}$$

Combining these two:

$$R^{\nu}_{t,s} = R^{M}_{t,s} \cdot \left(\frac{R^{\nu}_{t,T}}{R^{M}_{t,T}} \right)^{\left(\frac{s-t}{T-t} \right)}$$

So the beta version (not in the sense of an early version of a product, but instead using the sense of "beta" as a coefficient in the Capital Asset Pricing Model (CAPM)) of this would be as follows:

$$R^{\nu}_{t,s} = (\beta(R^{M}_{t,s} - 1) + 1) \cdot \left(\frac{R^{\nu}_{t,T}}{(\beta(R^{M}_{t,T} - 1) + 1)} \right)^{\left(\frac{s-t}{T-t} \right)}$$

Choose β such that the correlation of revisions with M is zero. This must be done jointly with the calibration of extrapolation parameters so we estimate the extrapolation and interpolation parameters in a manner consistent with each other.

2.3.2 Interpolation when one or both events have zero value

When the value at time T is equal to zero (which is the case for shutdowns, the outcome for roughly half of venture companies) we use an arithmetic method to interpolate value between the last funding event and the shutdown:

$$R^{v}_{t,s} = \left(\beta(R^{M}_{t,s} - 1) + 1) \right) \cdot \left[\frac{T-s}{T-t} \right]$$

For both types of interpolation, we choose β such that the correlation of revisions with the public market M is zero. This must be done jointly with the calibration of extrapolation parameters so that we estimate the extrapolation and interpolation parameters in a manner consistent with each other. Our current calibration results in a value of 1.37 for β.

2.4 Extrapolation for Unexited Companies

For unexited companies after their last known funding round, we estimate monthly value starting from the value at the company's last valuation event through the end of the Index. The estimation of these monthly values includes three components: a constant term, a market return multiplied times a coefficient, and a "decay parameter" multiplied times the number of months since the last funding round for the company. Specifically, the monthly returns after the final funding round for an unexited company are estimated using the following formula:

$$R^{\nu}_{s} = \alpha + \beta * R^{M}_{s} + \gamma * (s - t)$$

Where

R^{ν}_{s} is the monthly return for the company for month t

α is the constant term (derivation explained below)

β is the coefficient for the market return (derivation explained below)

R_s^M represents the monthly return for the ICB industry index, for the industry associated with that particular company

γ is the decay coefficient (derivation explained below)

$(s-t)$ is the number of months between the month for which the return is being calculated $(s)$and the month of the company's final finding round (t).

The extrapolation parameters $\alpha, \beta, and\ \gamma$ are derived from the historical data using a calibration technique that minimizes the sum of the squared differences between the monthly level of the full VC Index as currently estimated using all available data and the level as it would have been estimated using only the data available at a historical point in time. Specifically, the Index is estimated at quarterly intervals, with a one-quarter lag between the time when data become available and the first publication of the Index (e.g., for Q1 2009, we assume that the January, February, and March monthly Index levels are first published in June 2009, so that all of the funding and exit data for April, May, and June are also available in estimating the Index through March 2009). We also currently assume that data enters the DJVS database as it occurs—that is, there is no delay between the time a funding round or exit occurs and the time at which it becomes available for estimating the Index.

Using those assumptions, we initialize the parameters at a reasonable starting value and estimate the monthly Index levels as they would have been estimated for each quarter from 1998 through 2007. This represents a 10-year period, which would include 40 quarterly updates of the Index. We also calculate the "ultimate" Index levels for the same period using all available data through today. We then calculate the squared difference between the ultimate Index level and each initial monthly Index level from 1998 to 2007. We then adjust the values for $\alpha, \beta, and\ \gamma$ and run the exercise again, until we find the values that minimize the sum of the squared monthly differences. We plan to recalibrate these parameters on an annual basis, using a 10-year rolling window with a full two-year delay or lag between the end of the window and the "current" year (i.e., in 2011 the window will move from 1998–07 to 1999–08).

To simplify the calibration exercise, we limit our optimization exercise to use only integer values—thus, the final values must be scaled appropriately when they are applied to the final extrapolation formula as follows:

$$\alpha = \alpha_{(calibrated)} / 1{,}000{,}000$$
$$\beta = \beta_{(calibrated)} / 100$$
$$\gamma = \gamma_{(calibrated)} / 100{,}000$$

Using the 10-year period from 1998 to 2007, the current values for the extrapolation parameters $\alpha, \beta, and\ \gamma$are -13, 159, and -48 respectively. Thus, the current formula that is used to estimate the monthly return for each unexited company for each month after its final funding round is as follows:

$$R_s^v = (-0.000013) + 1.59 * R_s^M - 0.00048 * (t-s)$$

By applying these monthly return estimates to the value at the time of the company's final funding round, we are able to estimate a monthly dollar value for each unexited company through the end of the index period.

Please note that the estimated historical data that have been created with respect to the Index are back-tested. Estimated back-tested historical data represent calculations of how the Index might have performed in the past if it had existed. Back-tested performance information is purely hypothetical and is solely for informational purposes. Back-tested performance does not represent actual performance, and should not be interpreted as an indication of actual performance. Past performance is not indicative of future results. Index performance is not the same as fund performance, as it does not reflect management and other fees.

3. AGGREGATION

To calculate the value-weighted Index, for each month s the pre- and post-money value of included firms is summed over all included firms (denoted by i) by

$$Pre_s = \sum_{i=1}^{N} \nu_{s,i}$$

and

$$Post_s = \sum_{i=1}^{N} V_{s,i}$$

where ν represents the pre-money value and V represents the post-money value.

The Index return is calculated as

$$Ret_s = \frac{Pre_s}{Post_{s-1}}$$

The Index level (I) is set to 100 in January 1992. Thereafter

$$I_s = I_{s-1} Ret_s$$

Note that a firm is included in the calculation of Ret_s only if it exists in $Post_{s-1}$ and also in Pre_s, otherwise new firms and firms that leave the sample would change the computed returns of the Index. Note also that all pre- and post-money values above represent the value of the entire company, including both venture-funded and non-venture-funded portions.

APPENDIX A: DETAILED DISCUSSION OF INDEX CONSTRUCTION

This appendix provides additional detail for the calculation the Dow Jones Venture Capital Index (the "Index"). Specifically, it discusses the choice of estimation techniques for constructing the Index, alternatives that were considered, and why we believe the selected technique is superior.

As noted previously, the Index is built not from the fund-level return data used to construct other published venture capital indices, but rather from company-level pricing data, which we believe are necessary to create a timely, unbiased, monthly index of value that is analogous to traded-market indices such as the Dow Jones U.S. Total Stock Market Index. Building an index from company-level data faces two major challenges.

First, events that produce market values do not occur continuously for private companies as they do for traded stocks. Instead, they occur episodically, when a company raises new money, goes public, is acquired, or goes out of business. The Index includes only venture-funded companies that remain independent and private, so when companies go public, are acquired, or go out of business they exit the Index at the time, with the "exit event" providing the terminal value for that company's valuation stream. The episodic nature of pricing events requires that we estimate intermediate monthly values, both between pricing events (interpolated values) and for the months following the last funding round for those companies that have not yet exited (extrapolated values).

Second, the reporting of value by companies that complete funding rounds is voluntary. Often companies report that they completed a fund-raising but do not report the value at which shares were sold. The companies that do report values are not a random sample of all companies; rather, they are a biased sample—successful companies are more likely to share values, and funded companies that eventually go public are more likely to reveal their full funding history (albeit after the fact) as part of their initial public offering documents. Thus, it is important to account for and correct this bias when estimating non-revealed company values at funding rounds.

Finally, companies exit the Index in one of three ways: IPO, acquisition, or shutdown. For those companies that go public, the terminal value can be obtained from public documents, and for those that shut down, the terminal value is zero. For acquisitions, value of the company is not always reported. Sometimes an announcement indicates the value. Details of the acquisition can often (but not always) be learned from the acquiring company's public filings (10-Q, 10-K, or other). Higher values are more likely to be revealed, as they tend to generate more press coverage and general interest. Many acquisition values remain missing and must be estimated.

Where company values must be estimated, the following properties are desirable for estimates:

- unbiased (i.e., they will not consistently overestimate or underestimate the true value);
- minimum variance (estimate should minimize some measure of the difference between the estimated values and the true values); and
- realistic (for example, no negative values and no extreme changes from one month to the next).

The remaining sections below are divided as follows to provide additional information on the approach used for each type of estimate:

- company values at the time of a funding round,
- company values for acquisitions,
- monthly company values between valuation events, and
- monthly values for non-exited companies from the dates of their last funding rounds through the end of the Index (extrapolated values).

Estimating missing company values at the time of a funding round

As described, missing values for funding rounds are estimated using non-linear least squares (NLLS) on those companies with revealed values to estimate valuations for those funding rounds that are not revealed. The variables used in the regression are described in more detail in the White Paper. The equation used is as follows:

$$\begin{aligned} V_{pre} = exp(c &+ \beta_1 LogRaised + \beta_2 LogRtd + \beta_{3,4,5} Ind_{1,2,3} + \beta_{6,7,8,9,10} Bs_{1,2,3,4,5} \\ &+ \beta_{11,12,13} Fs_{1,2,3} + \beta_{14} BrPrior + \beta_{15} Unc + \beta_{16} Unc \cdot LogRaised \\ &+ \beta_{17} Unc \cdot LogRtd + \beta_{18} NotHasLastKV + \beta_{19} LogLastKV \\ &+ \beta_{20} LogYearsLastKV + \beta_{21} LogTSM + \beta_{22,23,...} Qtr_{2-T} + \epsilon) \end{aligned}$$

Though it is difficult to interpret the relationship to any specific deterministic variable from the NLLS results, we have examined the sign of the variables to determine if the basic relationship makes sense from an economic perspective. The following table presents the signs for the deterministic variables used in the regression to estimate non-revealed funding round values when all of the data currently available are used (1990 through present):

Intercept	-
log(Raised)	+
log(RTD)	+
PriorBridge	-
Uncensored	+
Uncensored*LogRaised	+
Uncensored*LogRTD	-
NotHasLastKV	+
LogLastKV	+
LogYearsLastKV	-
logWilshire	+
Industry Dummies (Omitted=IT)	
Health	-
Retail	-
Other	+
Business Stage Dummies (Omitted=Shipping)	
Startup	-
Development	-
Beta-testing	+
Clinical Trials	-
Profitable	+
Funding Stage Dummies (Omitted=Seed)	
Early	+
Late	+
Mezzanine	+
Intercept	-

Using the estimated coefficients $\widehat{\beta}$ obtained above using the revealed values in the historical data set, we then fill in estimated pre-money values $\widehat{V}_{pre}$ for all funding rounds in which this value is not revealed. This estimate, however, is biased because it represents an exponentiated value from the log model used to estimate the coefficients. To correct for this bias, we estimate a "scaling factor" (SF) using the value-weighted average difference between the actual revealed pre-money value V_{pre} and the estimated pre-money value $\widehat{V}_{pre}$ as follows:

$$SF = avg(V_{pre}) / avg(\widehat{V}_{pre})$$

The final pre-money value is then estimated as $(\widehat{V}_{pre} \cdot SF)$. After all missing pre-money values (V_{pre}) have been estimated, post-money values are set to $V_{post} = (V_{pre} + Raised)$, the pre-money value plus the amount raised in this round.

This estimation technique is superior for the following reasons:

As explained above, applying the scaling factor $avg(V_{pre}) / avg(\widehat{V}_{pre})$corrects for the bias inherent when using the exponentiated value from the log model, providing an unbiased estimation of the true value V_{pre}.

- By taking the exponential of the right-hand variable, we ensure that the resulting estimated value $\widehat{V}_{pre}$ is always positive.
- We estimate the pre-money value rather than the post-money value to eliminate the possibility of estimating a post-money value that is less than the amount raised in the round, which would result in a negative pre-money value. Using the approach above, we are assured of estimating positive values for both pre-money and post-money values for the firm at the time of the funding round.
- We estimate values for companies rather than returns because some company event streams will not provide a revealed value for any funding rounds (indeed, most companies that have recently entered the Index will have only one funding round thus far, and many of those funding rounds will not have a revealed value). If we were to estimate returns rather than values, we would still need some additional technique to estimate the initial value for many companies; this approach eliminates the need for such an additional step.

Estimating non-revealed company valuations for acquisitions

Values are also missing for nearly half of all acquisitions and must therefore be estimated. Acquisition values are somewhat different from funding rounds—no additional funds are raised. In addition, the reporting bias observed in funding rounds appears to be quite different from the reporting bias observed in acquisitions. For funding rounds, there are many motives for concealing value. For acquisitions, the main reason for concealment appears to be simply to avoid revealing less profitable outcomes.

To estimate missing acquisition values, we perform a nonlinear least squares regression similar to the one used to estimate non-revealed round values, though the form of the regression differs, using those explanatory variables shown to have the most explanatory power for estimating acquisition values. We also apply a scaling factor to address the bias resulting from converting from log values to level (dollar) values, estimated using the historical data on acquired companies. Finally, we adjust for the observed bias correlated with the difficulty of finding acquisition values as described above. This approach provides the same advantages as our approach to estimating non-revealed funding round values—we estimate only positive values for which any observed bias has been corrected.

Alternative Techniques for Estimating Firm Value

The standard techniques often used to value companies (especially publicly traded companies) cannot be applied here, as we seldom have information about company sales, revenues, income, cash flows, and so on. Alternatives to the regression approach we use could include the following:

Discounted Cash Flow Method

Estimating value by discounted cash flow requires an understanding of the size of the appropriate market for the company being analyzed, some estimation of the market share through time, and the net present value of the discounted cash flow. Such detailed information is simply not available for the companies included in the Index.

"Relative Value" or "Guideline" Method

This approach involves identifying a set of comparable publicly traded companies to the one being analyzed and using the standard financial price multiples (e.g., price-to-earnings or price-to-book value) for the firm being analyzed to estimate a value. Such an approach is inappropriate here because (1) most of the companies whose values are estimated would not correspond to publicly traded companies given their early development stage, and (2) earnings or book value information is unavailable or not comparable (e.g., the purpose of securing a VC investment is to complete the research and development necessary to produce the good or service, so that most early-stage VC-funded companies would not yet have any earnings).

Estimating monthly company values between two funding rounds or between a funding round and an exit event (interpolated values)

The interpolation of values between two valuation events (i.e., between funding rounds or between the final funding round and company exit) is necessary because the Index is published as a monthly series. When the two company values are both positive, the interpolated values are estimated as follows:

$$R^{v}_{t,s} = (\beta(R^{M}_{t,s} - 1) + 1) \cdot \left(\frac{R^{v}_{t,T}}{(\beta(R^{M}_{t,T} - 1) + 1)} \right)^{\left(\frac{s-t}{T-t}\right)}$$

When the second value is zero (i.e., a shutdown), the formula is as follows:

$$R^{v}_{t,s} = \left(\beta(R^{M}_{t,s} - 1) + 1) \right) \cdot \left[\frac{T-s}{T-t} \right]$$

Where

R = a monthly return,

s = the current month,

t = the month of the most recent valuation event,

T = the month of the following event (the next event after month s),

M = the Dow Jones Indexes based on ICB corresponding to the high-level industrial grouping for the firm,

β = the beta value that expresses the relationship between venture-backed companies and M, and

ν_i = the pre-money value for the firm.

Estimating the values between two known values is, on its face, a fairly straightforward proposition. One alternative to the approach used would be to carry the company forward at its last known value until the subsequent valuation event occurs; however, valuation events are often years apart, and this would often result in a large change in value in a single month and would not be a realistic reflection of the true valuation history for the company. Another alternative would be to construct a straight-line movement from the first valuation to the second. While simplistic, this would capture the movement in a reasonable fashion, but would fail to take into account irregular movements in value that likely occurred in the interim. By incorporating an industry-specific market component, the approach used to construct the Index provides a gradual value change with some variability commensurate with the performance of publicly traded companies in the same industry grouping as the company being estimated.

Estimating monthly values for each non-exited company from the time of its last funding round through the end of the Index (extrapolated values)

For each non-exited company included in the Index, we need to estimate monthly values from the time of its last funding round through the end of the index period, using only information that is currently available. To better understand the issues related to extrapolating values, consider the evolution of an individual company's value stream over time. A company will enter the Index when it has its first funding round. After that initial round, its value must be estimated via extrapolation until it has a follow-on round or exits, at which point the intervening monthly values will be estimated using the interpolation technique rather than the extrapolation technique. To provide liquidity for the VC investors, each funded company must eventually exit the Index via IPO, acquisition, or shutdown. Thus, Our goal with the extrapolation algorithm is to provide an unbiased, minimum-error estimate of the monthly company value that will eventually be either revealed via a pricing event or be interpolated between the penultimate and ultimate pricing events.

As explained in the White Paper, we estimate these values by calculating each monthly return using the following formula:

$$R_s^\nu = \alpha + \beta * R_s^M + \gamma * (s - t)$$

where R_t^M is the monthly return for the ICB industry index for the industry associated with that particular company and $(t-s)$ is the number of months between the month for which the return is being calculated (t) and the month of the company's final finding round (s). The extrapolation parameters $\alpha, \beta, and\ \gamma$ are derived from the historical data using a calibration technique that minimizes the sum of the squared differences between the monthly level of the full VC Index as currently estimated using all available data and the level as it would have been first estimated using only the data available at a historical point in time. For the calibration exercise, we have chosen to use a 10-year rolling window with a full two-year delay or lag between the end of the window and the "current" year (i.e., in 2010 the window is 1998–2007, and in 2011 the window will move to 1999–2008). This 10-year rolling window provides a sufficient period of time to conduct a rigorous calibration without relying too heavily on earlier historical data that may no longer be representative of the current relationships between the time since the company's earlier round, the public markets, and the current extrapolated company value.

This approach results in an unbiased estimate of company value that will result in the total Index most closely resembling the "ultimate" value that the Index will have in the future, when additional pricing events occur for the extant companies in the Index, and extrapolated values are replaced with interpolated values.

Alternative approaches we considered included the following:

- We considered a logit/probit model for each of the possible types of subsequent pricing events (funding round, IPO, acquisition, and shutdown) with an estimate of the associated value for each type of event based on the time passed since the company's last round, the total amount raised, public market performance, etc. This approach was fairly complicated to implement, and historical testing showed that the initial Index estimate generated from this approach was not as good a predictor of the "ultimate" Index levels as the approach we now use.
- We also tested a simple decay function, using historical company-level data to construct a simple function that used the time since last round and market movements to estimate value over time following the company's last round. While this provided a logical company-specific value estimate, it did not prove to be an unbiased, minimum-variance estimator for the index-level monthly returns. As such, it was discarded in favor of our current approach.

Including a public market component when interpolating or extrapolating firm values

The reason we use stock market data (specifically, the Dow Jones Indexes based on ICB corresponding to the high-level industrial grouping for the firm) when interpolating monthly firm values between valuation events and when extrapolating monthly firm values after the last known funding event is that we are confident that venture capital values and stock market values are highly correlated, and stock market data are available much more quickly than venture data are. Analysis of investors' returns on venture portfolios indicates that for the period prior to 2001, the beta for venture capital was roughly two (2.0) and the correlation of quarterly venture returns with stock market returns was about 0.8. For the period after 2001, the beta is one (1.0) and the correlation is still roughly 0.8. [2] As such, including a market component to the movement between funding events and for estimating monthly values for non-exited companies after their most recent funding event provides a more realistic estimate of the company's value.

Comparing the Dow Jones Venture Capital Index to Other Types of VC Indexes

Among the alternatives for tracking venture capital activity are the indices of the flows of funding into venture capital and various compilations of fund returns by other organizations. These index figures do not really measure the value of venture-funded companies, but only the flows of money to them. Funding and value are correlated, but they are not the same. Our goal is to provide an index of value.

Other indexes represent averages of returns on limited partnership funds. Our understanding is that the data for such indexes come from customers who provide the data for evaluations of fund performance. Our understanding is also that such indexes simply add the fund returns for each quarter and divide by the number of funds, so the returns are not value-weighted. We believe that these types of index returns are biased upward for the natural reason that customers do not need or want evaluations of the worst-performing funds. When funds perform poorly, their managers generally do not succeed in raising another fund. Data about them evaporates, and it is not possible to reconstruct it from any regularly gathered sources. Even though venture-funded companies are more numerous than venture capital funds, it is more feasible to construct complete funding and outcome histories from company data than from fund data.

---

[2] See Susan E. Woodward, "Measuring Risk for Venture Capital and Private Equity Portfolios," Sand Hill Econometrics Working Paper, 2009, available at http://papers.ssrn.com/sol3/papers.cfm?abstract_id=1458050, and Hall, Robert E. and Susan E. Woodward, "Benchmarking the Returns to Venture," Working Paper, available at http://papers.ssrn.com/sol3/papers.cfm?abstract_id=474181. Both of these works show the high correlation between venture returns and stock market returns.

We believe that nearly every venture-funded company becomes part of the DJVS database upon raising its first round of venture capital. Once in the database, DJVS attempts to follow the company through any additional funding rounds until it goes public, is acquired, or shuts down. Because the companies are systematically captured at first round, following them to their outcome is feasible. With significantly complete data on venture-funded companies, we have attempted to achieve an index of venture capital value that is value-weighted and continuously invested, which can be used to track returns to close to the entire U.S. venture capital market through time, and which can also be used to measure the risk of the U.S. venture capital market and its correlation of its returns with U.S. public stock market returns. As mentioned above, the compilations of fund returns, because they are often missing the worst outcomes, tend to produce return series that are biased upward and not value-weighted.

Sample of Venture-Funded Company Valuation Stream

To better illustrate the process by which the valuation stream for a single venture-funded company is established, consider the following example. First, the company's funding history available from DJVS is as follows:

Company ID	ICB Code	ICB Industry	Close Date	Event Type	Business Status	Pre-Money ($MM)	Raised ($MM)	Post-Money ($MM)
	9578	Technology	04/08/05	Early	Startup	6.00	6.00	12.00
XXXXXX	9578	Technology	08/01/06	Early	Product in Beta Test		15.00	
XXXXXX	9578	Technology	05/02/08	Early	Generating Revenue	55.00	12.00	67.00

This company is a technology firm that raised 33 million dollars over three funding rounds, the first in 2005 and the most recent in 2008. Value is revealed for two of those rounds. At this point the company has not had a liquidity event (i.e., it has not gone public, been acquired, or gone out of business), so it remains in the Index. The first step in constructing the full valuation stream for this company involves estimating a value for the funding round(s) where value is not revealed—here, only the second round value needs to be estimated. This is accomplished by applying the coefficients generated from the nonlinear least squares regression described above to estimate the non-revealed pre-money value, and the post-money value is calculated by adding the amount raised in the round to the pre-money value as follows (the estimated values here are shown in **blue**):

Company ID	ICB Code	ICB Industry	Close Date	Event Type	Business Status	Pre-Money ($MM)	Raised ($MM)	Post-Money ($MM)
XXXXXX	9578	Technology	04/08/05	Early	Startup	6.00	6.00	12.00
XXXXXX	9578	Technology	08/01/06	Early	Product in Beta Test	**35.64**	15.00	**50.64**
XXXXXX	9578	Technology	05/02/08	Early	Generating Revenue	55.00	12.00	67.00

The second step in constructing the valuation stream is to estimate a monthly value for each month between funding rounds. These interpolated values are estimated using the following formula:

$$R_{t,s}^{V} = (\beta(R_{t,s}^{M} - 1) + 1) \cdot \left(\frac{R_{t,T}^{V}}{(\beta(R_{t,T}^{M} - 1) + 1)} \right)^{\left(\frac{s-t}{T-t}\right)}$$

The table below shows the interpolated values for all of the months between the funding rounds for this company, with the interpolated estimates given in **green**. To illustrate how the calculation works, assuming a value of 1.37 for β, the interpolation formula would be as follows for 2005-05:

$$R_{t,s}^{\nu} = (1.37([5331.11/4806.01]-1)+1)\cdot\left(\frac{[35.64/12.00]}{(1(5710.35/4806.01]-1)+1)}\right)^{\left(\frac{1}{16}\right)} = 1.2131$$

Multiplying 1.1746 by the previous value of 12.00 results in the value of 14.56, as shown in the table for the Pre-Money value in 2005-05.

yyyy-mm	Month (s-t)	ICB Industry Index	Returns from Last Round	Pre-Money ($MM)	Raised ($MM)	Post-Money ($MM)
2005-04	-	4,806.01		6.00	6.00	12.00
2005-05	1	5,331.11	1.2131	**14.56**	-	**14.56**
2005-06	2	5,268.24	1.2601	**15.12**	-	**15.12**
2005-07	3	5,644.67	1.4556	**17.47**	-	**17.47**
2005-08	4	5,559.27	1.5058	**18.07**	-	**18.07**
2005-09	5	5,641.09	1.6193	**19.43**	-	**19.43**
2005-10	6	5,383.03	1.6071	**19.29**	-	**19.29**
2005-11	7	5,806.13	1.8714	**22.46**	-	**22.46**
2005-12	8	5,697.68	1.9271	**23.13**	-	**23.13**
2006-01	9	6,010.51	2.1780	**26.14**	-	**26.14**
2006-02	10	6,015.23	2.3005	**27.61**	-	**27.61**
2006-03	11	6,144.49	2.4939	**29.93**	-	**29.93**
2006-04	12	6,152.82	2.6360	**31.63**	-	**31.63**
2006-05	13	5,651.54	2.4942	**29.93**	-	**29.93**
2006-06	14	5,472.77	2.5237	**30.28**	-	**30.28**
2006-07	15	5,206.20	2.4929	**29.91**	-	**29.91**
2006-08	-	5,710.35		**35.64**	15.00	**50.64**
2006-09	1	5,885.95	1.0355	**52.43**	-	**52.43**
2006-10	2	5,987.32	1.0529	**53.32**	-	**53.32**
2006-11	3	6,266.60	1.1120	**56.31**	-	**56.31**
2006-12	4	6,163.79	1.0809	**54.73**	-	**54.73**
2007-01	5	6,216.48	1.0862	**55.00**	-	**55.00**
2007-02	6	6,139.36	1.0615	**53.75**	-	**53.75**
2007-03	7	6,125.83	1.0516	**53.25**	-	**53.25**
2007-04	8	6,450.77	1.1190	**56.66**	-	**56.66**
2007-05	9	6,709.82	1.1706	**59.28**	-	**59.28**
2007-06	10	6,858.06	1.1965	**60.59**	-	**60.59**

yyyy-mm	Month (s-t)	ICB Industry Index	Returns from Last Round	Pre-Money ($MM)	Raised ($MM)	Post-Money ($MM)
2007-07	11	6,887.36	1.1955	**60.53**	-	**60.53**
2007-08	12	7,125.98	1.2409	**62.83**	-	**62.83**
2007-09	13	7,367.70	1.2864	**65.14**	-	**65.14**
2007-10	14	7,631.22	1.3360	**67.65**	-	**67.65**
2007-11	15	6,999.36	1.1898	**60.25**	-	**60.25**
2007-12	16	7,009.93	1.1845	**59.98**	-	**59.98**
2008-01	17	5,958.28	0.9506	**48.13**	-	**48.13**
2008-02	18	5,847.00	0.9207	**46.62**	-	**46.62**
2008-03	19	5,867.22	0.9192	**46.54**	-	**46.54**
2008-04	20	6,290.53	1.0027	**50.77**	-	**50.77**
2008-05	-	6,718.61		55.00	12.00	67.00

The final step in completing the full valuation history for this company is to estimate monthly values from the time of its last round through the end of the index. For this example, assume the index ends in December 2009. As explained above, the formula used to estimate extrapolated values is as follows:

$$R_s^v = \alpha + \beta * R_s^M + \gamma * (s - t)$$

As discussed in the White Paper, the values currently calibrated for $\alpha, \beta, and\ \gamma$ are -0.000013, 1.36, and -0.00034, respectively. The table below shows the extrapolated values for our sample company. To further illustrate how these values are estimated, consider the first extrapolation value for 2008-06:

$$R_s^v = -0.000013 + 1.59 * ((6015.34 / 6718.61) - 1) + (-0.00048) * (1) = -0.1669$$

Applying this return times the previous value: $67.00 * (-0.1669 + 1) = 55.82$. The table below shows the values using the extrapolation formula in **red**.

yyyy-mm	Month (s-t)	ICB Industry Index	Returns from Last Round	Pre-Money ($MM)	Raised ($MM)	Post-Money ($MM)
2008-05	-	6,718.61		55.00	12.00	67.00
2008-06	1	6,015.34	-0.1427	**55.82**	-	**55.82**
2008-07	2	5,962.85	-0.0126	**54.99**	-	**54.99**
2008-08	3	6,144.83	0.0405	**57.58**	-	**57.58**
2008-09	4	5,103.32	-0.2319	**41.95**	-	**41.95**
2008-10	5	4,226.97	-0.2353	**30.39**	-	**30.39**
2008-11	6	3,679.48	-0.1782	**24.05**	-	**24.05**
2008-12	7	3,759.83	0.0273	**24.80**	-	**24.80**

yyyy-mm	Month (s-t)	ICB Industry Index	Returns from Last Round	Pre-Money ($MM)	Raised ($MM)	Post-Money ($MM)
2009-01	8	3,632.54	-0.0488	**23.37**	-	**23.37**
2009-02	9	3,445.60	-0.0731	**21.36**	-	**21.36**
2009-03	10	3,956.38	0.1982	**26.29**	-	**26.29**
2009-04	11	4,528.14	0.1928	**32.19**	-	**32.19**
2009-05	12	4,627.40	0.0257	**33.12**	-	**33.12**
2009-06	13	4,842.32	0.0587	**35.36**	-	**35.36**
2009-07	14	5,432.44	0.1610	**41.98**	-	**41.98**
2009-08	15	5,555.73	0.0258	**43.19**	-	**43.19**
2009-09	16	5,844.62	0.0653	**46.43**	-	**46.43**
2009-10	17	5,619.71	-0.0581	**43.21**	-	**43.21**
2009-11	18	5,881.46	0.0572	**46.03**	-	**46.03**
2009-12	19	6,307.10	0.0920	**50.91**	-	**50.91**

APPENDIX B: DATA PROCESSING PRIOR TO ESTIMATION OF INDEX

The data as received from DJVS must first be processed before it can be used to generate the VC Index. The following steps summarize the main processing activities:

1. We do not include every firm in the DJVS data in the Index. Excluded companies are those which we believe are not best described as venture-backed companies. Typically they are very large and only a small proportion of their funding comes from venture sources. We remove these firms entirely from the data, and they are not included in any part of the Index calculation.

2. Any records that are missing a CloseDate value are removed (all events must have a date to be included in the Index estimation).

3. The DJVS data include event types that are not appropriate for the purposes of generating the VC Index, such as PIPEs, buyout rounds, etc. These events are removed.

4. Additional historical IPO information as researched by Sand Hill Econometrics is integrated into the data set. Specifically, the final value for a company in the Index that exits via an IPO must be equal to the total number of shares outstanding prior to the IPO times the IPO offering price. This information has not historically been captured by DJVS—rather, the DJVS data captures the fully diluted value of the company, including any additional shares issued in the IPO.

5. Additional "uncensored" valuations as captured by Sand Hill Econometrics are integrated into the DJVS data. The Sand Hill Database of venture-backed companies included information that came from uncensored sources—specifically, quarterly or annual reports provided by the VC fund general partners to the limited partners (investors). These data include both rounds and exits, and represent information that is sometimes neither revealed publicly nor provided to DJVS. Historical research has shown that values that are publicly shared are on average higher than values not shared—that is, the general partners and companies themselves are less likely to disseminate information on funding rounds or exits when those events reveal a low company value, etc. An essential part of the Index calculation includes correcting for this reporting bias, so the uncensored data available from the Sand Hill database is integrated into the DJVS data and marked as uncensored if the DJVS data did not have a revealed value for the event.

6. The "recap" and "restart" companies are processed. Both of these designations imply that the prior venture investors are wiped out and new investors "start over" with the company. Prior to October 2009, a "RECAP" in the DJVS data indicated that the company was taken over by buyout investors, and a "RESTART" indicated that the new investors were VC investors. As of October 2009, however, these designations were reversed (RECAP indicates new VC investors, RESTART indicates buyout investors). Thus, for companies with a RECAP we create a shutdown record equal to the recap date, then assign a new company ID for the events starting with the RECAP round to treat those rounds (and, potentially, exits) as belonging to a new company. For companies with a "RESTART" round, we simply create a shutdown record equal to the date of the RESTART and ignore the events after that point.

7. If a company has more than one round that occurs within the same month, the raised amounts are combined and the maximum postval value is used to create a single financing round for the company for that month.

Treatment of Outliers

There are a few VC-backed companies which are orders of magnitude more valuable than the majority of the sample. Although we consider these outlier firms to be valid venture-funded firms and we include them in the Index calculation, we do not believe they are helpful in inferring value for the companies that do not reveal value, simply because values so large would not be feasible to conceal. Their inclusion would have a misleading influence on the estimated values for companies not revealing value. Thus we exclude from the estimating regression any financing round that has a revealed value more than four standard deviations from the mean for any 10-year rolling period.

APPENDIX C: CALIBRATION OF ACQUISITION MISSING VALUE ADJUSTMENT FACTOR

We model the relationship between the value of an acquired firm as an exponentially decreasing function of the amount of difficulty involved in obtaining that value. We use four datasets to model the difficulty level: Source A (an M&A database), DJVS data, Source B (another M&A dataset Sand Hill has access to), and the merged DJVS+Sand Hill database. We let x, the amount of difficulty involved in obtaining acquisition value, be proxied by the percentage of acquisition values each dataset contains.

We model the average revealed value of an acquired firm obtainable with difficulty x as

$$A_r(x) = \frac{1}{x}\int_0^x V_0 \cdot e^{-\alpha s} \mathrm{d}s = \frac{V_0(1-e^{-\alpha x})}{\alpha x}$$

We calibrate α over the following data, letting the mean revealed value $(M_{r,i})$ be our estimate of $A_r(x_i)$

Source (i)	Percentage Found (x)	Mean Revealed Value $(M_{r,i})$
Source A	16%	$181 million
DJVS	41%	$143 million
Source B	50%	$120 million
DJVS+SH	56%	$94 million

The choice of α determines the value of the parameter $V_{0,i}$ for each dataset, i, by

$$V_{0,i} = \frac{\alpha x_i M_{r,i}}{1-e^{-\alpha x_i}}$$

In principle, $V_{0,i}$ should be the same for all observations, so we vary α to minimize the coefficient of variation of $V_{0,i}$ over these four observations. That is, we minimize its standard deviation divided by its mean. This computation yields a value of α=3.7.

We can calculate the value of unreported acquisitions corresponding to difficulty x_i (that is, the value of acquisitions too difficult to get if we are expending only up to effort x_i) using the integration equation for $A(\cdot)$ for firms between difficulty x and 1. Since this estimates the average value of non-revealed firms, we denote it by $A_u(x)$.

$$A_u(x) = \frac{1}{1-x}\int_x^1 V_0 \cdot e^{-\alpha s} \mathrm{d}s = \frac{V_0(e^{-\alpha} - e^{-\alpha x})}{\alpha x - \alpha}$$

For the DJVS dataset this value is $24.78 million. Since the average value of reported acquisitions is $143 million, the average unreported acquisition is worth 24.78/143=.1733 as much as the average reported acquisition.

To calculate the adjustment factor, we calculate the mean value of the reported firms (M_r) and the estimated mean

value of non-revealed acquisition values using the regression equation described in the acquisition section before applying any adjustment. Then

$$\lambda = \frac{.1733 \cdot M_r}{A_u}$$

Where M_r and A_u refer to the calculated values corresponding to DJVS. Filling in for the above equation results in a value of about 0.20 for λ.

The adjustment factor depends on the target ratio of mean non-revealed to mean revealed acquisition values. This, in turn, depends on the dataset we are examining. The Index going forward will use only DJVS data, whereas the data we use to calculate the historical Index uses DJVS+SH data. This implies that, going forward, we may need to recalculate the value of λ periodically, just as we need to recalibrate other parameters used to estimate non-revealed company values.