

THOMSON REUTERS INDICES

VENTURE CAPITAL RESEARCH INDEX METHODOLOGY

Thomson Reuters Venture Capital Research Index Methodology

Version 1.0

Authors: Grant Farnsworth¹, Jeff Kenyon², Florian Fischer³

Contents

Index Description.....	3
Data Sources	3
Universe Identification	3
Computing Missing Acquisition Valuations	4
Computing Missing Round Valuations	4
Generating Monthly History for each Firm.....	4
Portfolio Creation.....	5
Rebalance Periods and Index History.....	5
Detail on Regressions	5
Estimation of Missing Firm Values	5
Estimation of Firm Value at a Funding Round when Value is Not Revealed.....	5
Firm and Round Characteristic Variables.....	5
Macroeconomic Variables.....	6
Estimation of Firm Value for Acquisitions when Value is Not Revealed.....	6
Monthly Interpolation Between Valuation Events.....	7
Interpolation when the firm has a value greater than zero at both events	7
Interpolation when one or both events have zero value	8
Extrapolation After Last Funding Round.....	9
Aggregation.....	10
For Further Information	10

¹ DSC Quantitative Group

² Thomson Reuters

³ Thomson Reuters

Index Description

The **Thomson Reuters Venture Capital Research Index** (“TR VC Research Index”) tracks the gross performance of the US venture capital industry through a comprehensive aggregation of venture-funded private company values. The TR VC Research Index is market capitalization weighted and published quarterly.

The final methodology for the index was developed through a partnership of Thomson Reuters, DSC Quantitative Group, and academic advisors.

Thomson Reuters publishes the index data on a quarterly basis based on Thomson Reuters Private Equity Performance Database analyzing the cash flows and returns for over 1,335 US venture capital partnerships.

Interested readers will find the Thomson Reuters document *Standard Private Equity Data Feed* to be useful in understanding the primary source of private equity data available for the project. No effort is made to replicate these documents here, or incorporate their content.

Data Sources

The Thomson Reuters Standard Private Equity Data Feed (hereafter referred to as PE) defines the universe of companies and round events. Valuation data for venture backed exits from Thomson Reuters Deals content is also used and merged with the data set.

Market and Sector Indices Data is used in both regressions, as well as in the interpolation/extrapolation. Sectors used are mainly high level “economic” sectors, with the exceptions of Healthcare and Technology:

- Basic Materials
- Cyclical Consumer Goods and Services
- Energy
- Financials
- Technology Equipment
- Software
- Industrials
- Noncyclical Consumer Goods and Services
- Telecommunications
- Utilities
- Healthcare Services
- Pharmaceuticals

Universe Identification

The starting universe of the firms of interest in the venture capital universe is defined by the coverage of the Thomson Reuters Private Equity Datafeed:

Data is retrieved, duplicates are removed, and invalid entries are eliminated.

The data is then complemented with the computed columns, derived from these existing round data.

The disclosed valuation provided in the feed is post-money, so the pre-money value is the disclosed valuation (if any), minus the round total. Note that if round total is greater than disclosed valuation, pre-money valuation is set to zero.

For each firm with more than one round record, summary values are computed and a market value as of the round date is added.

Exit data is added to the single data structure. “Secondary Sales” are removed, as these are sales between venture capital companies, leaving IPOs, Trade Sales, and Buybacks.

For each ID in the round data, we identify the first exit round as the exit of interest. It is possible for a firm to be taken public, return to private status, and then exit from private status a second time, but events following the first exit are not considered in the index.

For IPO exits, we use a Post Offer Value that adds the Proceeds Amount and the Overallotment Sold All Markets as the valuation.

For the index launch in October 2012 with a history from 1996, this universe selection results in a sample of over 22,000 venture capital companies.

Computing Missing Acquisition Valuations

Regressions are done across the entire back-test period. This means that the body of data used in the regressions is increasing with each rebalance, rather than using a rolling data window of fixed width. It also means that monthly index values may shift as new data are incorporated into the regressions.⁴

Computing Missing Round Valuations

A “Heckman correction” is used to fill in missing round valuations.

The Heckman correction method introduces the inverse Mills ratio as an additional explanatory variable into the regression model.

The Heckman correction method is described in more detail in the document *A Selection-Bias Corrected Method for Filling Missing Values*.⁵

Generating Monthly History for each Firm

Prior to creating the monthly series, a synthetic table of “failure events” is created. Because firms are not explicitly declared as bankrupt in the data, we assume that if a firm’s *current* status in the PE data source is “defunct,” then the firm failed one year after the latest funding round. If we have received no data about any events associated with the firm for more than five years, we assume it failed five years after the last funding round even if the current status has not yet been changed to “defunct” in the PE data source.

The round and exit data, with missing valuations filled, are then used to create a regular monthly data series for each firm, since funding rounds or exits may come months or years apart.

⁴ In order to take advantage of all information, a window of 8 recent quarters is restated on a quarterly basis should the index valuation change due to newly available data.

⁵ Details available on request

Portfolio Creation

The last step of the process is the computation of the monthly index values. For each rebalance period, we:

1. Determine the potential candidates for the previous and current months.
2. Determine the index return, which is the ratio of the sum of the *pre*-money valuations for the candidate set in the current month, over the sum of the *post*-money valuations for the candidate set for the previous month.
3. Determine the sector weights for the month.

Rebalance Periods and Index History

Index Inception Date: Base value of 100 as of 30 December 1994. Prices are computed quarterly, as of the last business day of the month, beginning on 31 January 1995.

Rebalance Date: The Research Index is rebalanced and recomputed on a quarterly basis.

Detail on Regressions

The index uses valuations implied by financing rounds along with other firm characteristics to construct a monthly value-weighted index.

Estimation of Missing Firm Values

Firm value at the time of an IPO 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 acquisition events 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.

Estimation of Firm Value at a Funding Round when Value is Not Revealed

We use the information available about firms that have revealed values during funding rounds to estimate the value of firms for which the value is not revealed. We use the following variables as predictors of value.

Firm and Round Characteristic Variables

- the log of the amount raised in this round (*LogRaised*)
- the log of the cumulative amount of money raised excluding the current round (*LogRtd*)
- industry indicators: Software, Technology Equipment, Healthcare,... (*Ind_{1,2,3,...}*)
- business stage indicators: Profitable, Shipping, Clinical,... (*Bs_{1,2,3,...}*)
- funding stage indicators: Expansion, Later Stage,... (*F_{1,2,...}*)
- indicator for whether the immediately prior event for the company was a bridge loan (*BrPrior*)
- indicator denoting if there is no previous funding round with a revealed (known) value (*NotHasLastKV*)
- the log of the last revealed (known) value, if available (*LogLastKV*)
- the log of the time elapsed (in years) since the last revealed (known) value, if available (*LogYearsLastKV*)

Macroeconomic Variables

- the log of the Thomson Reuters United States Total Return Index level ($LogTSM$)
- indicator for each quarter (Qtr_t)

These variables were selected using the following criteria:

- the variables must be regularly available (i.e., though some data such as income or balance sheet information would be very useful as predictors of non-revealed values, such information is simply not available for these venture-funded companies); and
- initial tests using OLS indicated those variables that were statistically reliably related to value.

Using only the (non-exit) funding rounds that have a pre-money value, we estimate the following regression using nonlinear least squares (NLLS):

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

where c denotes an intercept term and ϵ is the error term.

Note that the coefficients corresponding to multiple variables (such as $Bs_{1,2,3,\dots}$) represent a set of distinct coefficients. Similarly, Qtr_{2-T} denotes a set of dummy variables, one for each quarter since the beginning of the index, omitting the first quarter. Thus $\beta_{22,23,\dots}$ is of dimension $1*(T-1)$.

Using the estimated coefficients $\hat{\beta}$ obtained above, we calculate estimated pre-money values \widehat{V}_{pre} for all funding rounds for which this value is missing.

Estimation of Firm Value for Acquisitions when Value is Not Revealed

To estimate the value of firms for which acquisition value is missing we perform a least squares regression similar to the one used to estimate non-revealed round values. The form of the regression used to estimate non-revealed acquisition values is as follows:

$$\begin{aligned} V_{acq} = \exp(& c + \beta_1 TSM + \beta_2 TSM^2 + \beta_3 TSM^3 + \beta_4 LogTSM + \beta_5 LogRTD + \beta_6 LogLastKV \\ & + \beta_7 NotHasLastKV + \beta_8 ElapFirst + \beta_9 ElapFirst^2 + \beta_{10} ElapFirst^3 \\ & + \beta_{11} ElapLast + \beta_{12} ElapLast^2 + \beta_{13} ElapLast^3 + \beta_{14} BrPrior + \beta_{15} LogMonth \\ & + \beta_{16} Info + \beta_{17} LastKVGr250 + \epsilon) \end{aligned}$$

where

TSM is the level of the Thomson Reuters US Total Return Index

$LogTSM$ is the log of the level of the Thomson Reuters US Total Return Index

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

$LogLastKV$ 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
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
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 company was greater than \$250 million

We eliminate extreme outliers from this regression, this time by removing any company that was acquired with a revealed value of USD 400 million or more. The remaining acquisition data with revealed values are then used in the above regression to estimate parameter coefficients which can be used to estimate values for those acquisitions that do not have revealed values.

We address the bias that results from converting from log values to level (dollar) values by estimating a "scaling factor" (*SFacq*) 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 difference, we construct an adjustment (downward) to the estimated missing acquisition values. We calibrate this adjustment factor λ and multiply the estimated value by this adjustment factor to get the final estimated value: $\widehat{V}_{acq} * \lambda$.

For a detailed discussion of how this adjustment factor is calculated see Appendix C.

Monthly Interpolation Between Valuation Events

After estimating missing values we have a flat file containing a row for each valuation event, including a value, either observed or estimated. We seek to generate a monthly pre- and post-money value for each firm in the sample between its first and last known valuation event.

In months containing a valuation event, the pre- and post-money values are known (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 Thomson Reuters Business Classification determined sector index level (corresponding to the sector classification for the firm in question) in month i by M_i , the post-money value by V_i and the pre-money value by v_i .

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:

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

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 re-write the equations above to use returns rather than levels.

Let

$$R_{t,s}^v = v_s / V_t$$

Now we rewrite the above formula for γ as

$$\gamma = \left(\frac{\log\left(\frac{R_{t,T}^v}{R_{t,T}^M}\right)}{T-t} \right)$$

and

$$R_{t,s}^v = R_{t,s}^M e^{\gamma(s-t)}$$

Combining these two:

$$R_{t,s}^v = R_{t,s}^M \left(\frac{R_{t,T}^v}{R_{t,T}^M} \right)^{\frac{(s-t)}{(T-t)}}$$

So the beta version of this would be as follows:

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

The parameter β is taken from the calibration described in the section on extrapolation.

Interpolation when one or both events have zero value

When either the pre-money value at time T or pre-money value at t is equal to zero (which can happen for shutdowns) we use an arithmetic interpolation method:

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

Our current calibration results in a value of 1.195972 for β .

Extrapolation After Last Funding Round

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

$$r_s^v = \alpha + \beta * r_s^M$$

where

r_s^v 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 Thomson Reuters sector index, for the industry associated with that particular company

Note that the use of lower case r for returns in these expressions denotes the use of log returns. For example, $r_s^M = \ln(1 + R_s^M)$, where R_s^M is the simple return on portfolio M in month s.

The extrapolation parameters α and β are derived from the historical data using a calibration technique that minimizes the sum of the squared differences between realized (observed) historical firm valuations and their predicted valuations using the extrapolation function. That is, for every pair of observed firm events, we predict the value at the second event using the extrapolation function and compare that value to what was observed at the time of the second event. We choose α and β by minimizing the sum of squared differences over all valuations.

Performing this calibration over observed data from January, 1995 to October 2007 (we exclude recent periods because they are still subject to revision) we compute the following:

$$\alpha = -0.0122633$$

$$\beta = 1.195972$$

Notice that the negative coefficient on α implies a monthly decay, which accounts for the fact that firms with delayed rounds or delayed reporting tend to have worse returns than those with frequent events and immediately reported returns.

Thus the extrapolation relationship is:

$$r_s^v = -0.0122633 + 1.195972r_s^M$$

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.

$$V_s = V_{s-1} \exp(r_s^V)$$

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:

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

and

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

where v represents the pre-round value and V represents the post-round value.

The index return is calculated as:

$$Ret_s = \left(\frac{Pre_s}{Post_{s-1}} \right)$$

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

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

Note that a firm is only included in the calculation of Ret_s if it exists in $Post_{s-1}$ and 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.

For Further Information

Visit the Thomson Reuters Indices [website](#) or contact Index_Queries@thomsonreuters.com.