

## Benchmarking

Benchmarks for investing serve many purposes. At the most basic level, they serve to inform investors about the change in value of some group of assets so that they have some sense of how markets are changing without reviewing many individual stocks. They also serve to establish an average performance against which investors can compare their own results to determine whether they are doing better or worse than average. Once comparison against the average is accomplished, we can expect the benchmark to be used to get at the causes of performance results. For example, a return above the benchmark could be due to greater risk, to genuine superior performance, or to avoidance of particular subsets of the asset class.

### Stock Market Benchmarks

One of the most followed benchmarks for the US stock market, the Dow Jones Industrial Average, goes back to the last decades of the 19<sup>th</sup> century. The Dow was designed to fulfill only the first purpose above: to be a summary number for the changes in stock values overall. The original DJIA was simply the sum of the prices of a group of well-known companies. As time passed, some of these companies split their shares, or did reverse splits, and as they did the formula for computing the index changed. Despite its primitive nature, it was found useful by many as an indicator of market-wide changes.

As performance evaluation grew more scientific and quantitative folks got involved, the shortcomings of the Dow were recognized and new indices were

created, notably the Standard & Poor's 500. S&P's index, like the old DJIA, is not an index of the entire market, but instead is 500 companies chosen by S&P to represent the market. The S&P500 is value-weighted, which means that the returns of more valuable companies have a much bigger impact on the index return than do the returns on small companies. The weights are strictly proportional to the value of the company's total outstanding stock. Stock splits and stock dividends have no impact on the index because it's only the total value of the company that matters, not the value of an individual share. The S&P produces two versions of its index, one whose returns include cash dividends (the most used, most familiar) and another that includes only changes in value or capital gains. The returns on the S&P are representative not just for the 500 companies included, but for a much broader portfolio. In other words, S&P did and does quite a good job of selecting 500 companies that are representative of the market overall.

Performance analysis continued to become even more scientific and professional, and this created interest in an even more complete index, which we have now in the Wilshire5000. The Wilshire attempts to include all public companies in its measure of changes in market value. "Public" means a company listed on an exchange and reporting to the SEC. The set of companies that is public on any given day is more difficult to track than it might seem, because tiny companies are de-listed and re-listed on exchanges with considerable frequency. For example, if a small company fails to file its financial reports with the SEC, it will likely be de-listed and moved from the regular

Nasdaq market to the OTC market or pink sheets. This does not mean it ceases to exist or that its stock ceases to trade. It just does not trade in the “listed” market.

A quality investors often say they want in an index is for it to be “investable”. This gives the old DJIA some points: all of its companies are big established companies and you can simply buy the requisite number of shares in each. The S&P500 is also investable in the sense that you can buy the 500 stocks that are in the S&P. But can you buy them at the *value* of the S&P? The answer is: Not quite.

The reason the answer is “not quite” is simple. Many of the companies in the S&P are traded in very thick markets, with prices printed from moment to moment. On any day, their closing prices will be for trades at the moment the market closes. But others are fairly thinly traded, so the last price for these thinly traded stocks may not be a trade at market close. If it is not, then the price at which you would buy the stock now if you tried is likely not the price used to compute the “closing” value of the S&P. You can buy the stocks, but you cannot necessarily buy them at the value of the S&P because some of the prices in the S&P are a few hours or possibly a few days stale.

Of course, we now have exchange-traded funds and future contracts in the S&P, and if you buy the S&P through these, then the index is absolutely investable. The small departures of the S&P index from the values seen on the ETFs come from the staleness of a few prices in the S&P.

Staleness is present in the Wilshire5000 to a somewhat greater degree because the Wilshire goes deeper into the smaller, thinly traded companies. There are more small companies without a current price. On the other hand, the Wilshire is more representative of the market as a whole because it aims more precisely at the whole market, and it exhibits a slightly greater variance over time than does the S&P because it includes these smaller companies. The composition of the Wilshire changes much more often than does the composition of the S&P (which only changes when S&P decides it needs to be changed – expect Google to be added soon) because of the frequent entry and exit of companies from the public markets. But because of difficulty of getting current prices for all of the companies, it is even harder to buy the Wilshire5000 than to buy the S&P at the published values.

For risk measurement and performance analysis and diagnosis, both indices are useful, but the Wilshire will capture more of the variation in the value of small companies than will the S&P. For large well-diversified portfolios, the answers they will give to questions about performance and its sources will be very similar. For portfolios focusing on smaller companies, the Wilshire will be the more informative.

## **Benchmarks for Alternative Assets**

There is an important and growing class of investments known as “alternative assets”. This includes hedge funds, venture capital funds, buyout funds, as well as funds that invest in real estate, oil and gas, distressed debt, and other not-quite-so-mainstream assets. All are organized as limited partnerships. These partnerships hold at least some assets that are not traded in any public market. None is regulated by the SEC.

Because the assets of these partnerships are not traded, there is no place to observe their prices, so the stale pricing phenomenon in private equity values and valuations is *much* more acute than staleness in the public markets. Value reporting is done by the general partner, who reports to the limited partners at most once per quarter. For venture capital, the values reported are those as of each company’s most recent fund raising. This value could be a few weeks old or a few years old. Buyout deals are reported at values estimated by the general partner based on stock market comparables (since the companies in a buyout portfolio usually are more mature companies comparable to others in the stock market). For buyouts also, valuations will often be stale because the comparables are not absolutely up-to-the-minute.

Given the extreme nature of the stale pricing problem for “alternative” assets, building a benchmark for these investments is a challenge.

**Two approaches can be taken to constructing such a benchmark. One begins with returns on investment**

**funds, and the other begins with values of the assets themselves.** The fund return approach is used by both Venture Economics and Cambridge Associates. With this approach, data is collected from investors on the returns in limited partnerships. If it were possible to get data on every limited partnership, in principle all of the portfolio companies would be represented too. In principle, it should be possible by value-weighting the partnership returns to construct an index from this data that mimics the results of investing in all of the funds, which should be the same as investing in all of the companies.

However, if the return data are not complete (some funds are not represented), the average returns across the partnerships will be different from the changes in value of the set of companies. In practice, the CA and VE partnership return benchmarks are neither complete nor value-weighted. They take whatever returns they can obtain and simply add them up and divide by the number of partnerships.

**Even if the return data were complete and assembled into a value-weighted return series, the “returns” will not be comparable to stock market returns.**

This is because while stock market index returns have staleness measured in hours or days at most, because valuations are mainly quite close to current, alternative asset quarterly returns are constructed from prices that are months or years old. To compare venture fund returns with public market returns is thus to compare an entirely current return (stock market) with returns that are a mix of returns over different intervals of time (the venture returns).

**Another approach is to use data from the asset level instead of the fund level to build the index.** For venture capital, this means using company values. Venture-funded companies typically do a round of funding once every year to two years. As a result, if each company's value is carried at its last known price, an index built from raw company values will be very much like an index from fund returns, and it will exhibit profound staleness. It will have the advantage of incorporating all of the relevant assets, but it will value them with mainly stale prices.

Thus, yet another approach is a statistical one, which is to take all of the company valuation data, and use the information from current transactions to value all of the companies represented in the index. **This is the only approach that would in principle produce an index that incorporates all of the relevant companies and prices them at current values.** Inspiration for this approach comes from the work on building indices for the value of single-family houses that began in earnest in the late 1980s. The statistical approach has been very successful in building house price indices, and these new indices are more informative than were the old indices built from appraisals. After all, appraisals are valuations that are filtered through human brains based on recently observed transactions. The statistical approach is also based on observed transactions, but treats the data in a systematic and rigorous way. Statistically-built indices are much more successful at predicting actual transaction prices than appraisal-based indices were.

Real estate investments, which are also organized as limited partnerships (and hold mainly commercial properties, not residential) are subject to the same valuation issues. The Pension Real Estate Association through the National Council of Real Estate Investment Fiduciaries (NCREIF) has sponsored research in the building of benchmarks. **Their approach has been to build both kinds of indices: one from returns from partnerships, and another based on property values.** It is useful to compare and contrast the approaches, because the issues for real estate are the same as for venture capital, buyouts, and other illiquid investments.

With the universe of partnership return data, plus information on the total invested in each partnership, one could construct the values and returns resulting from investing in all of the companies (or properties) in all of the portfolios. The totaled portfolio results will be equal to the summed company results (minus the fees to the portfolio managers). With both approaches, the valuations will at any point be a combination of current and past (stale) valuations.

To bring the valuations current, there is no avoiding statistics. The marking-to-market can be done three different ways:

1. Use statistical methods to estimate value changes from one date to the next for every company. This can be done via classic repeat-sales estimation, or done the way we build the sand hill index, by estimating directly a value for every company at every date. This approach calculates a current value for every company, which is the marked-to-market value.

2. Use statistical methods to compare part-stale portfolio returns to returns on a portfolio with no (or almost no) staleness (such as the stock market), recover the pattern of staleness in portfolio returns, and use the estimated staleness to mark the portfolio to market. This is a regression approach. The regression does several jobs at once. First, it measures how the current returns on the (part stale valued) portfolio are related to returns on the stock market, both current and lagging. The regression coefficients also reveal what fraction of the portfolio is one period stale, two periods stale, etc. The regression also measures the true current relationship between the portfolio and the stock market index, which is just the sum of the coefficients on the set of current and lagging returns. The actual stock market returns plus the estimated relationship between the stock market and the portfolio is then used to bring the stale parts of the portfolio current.

3. Exploit the property that “natural” returns are not serially correlated. The part-stale return series will exhibit serial correlation, so the approach is to estimate the serial correlation in the part-stale return series, and extract the non-serially correlated core (the “signal”, or the new element from each return) from the series. This is also referred to as generating a new, uncorrelated series using a Kalman filter, which filters out the part of return that is stale.

Approach number one will, in general, likely be the most accurate. It is also the most laborious.

To see why the first is potentially the most accurate, let’s start with the third and think about what the earlier approaches add.

The third approach uses only information about the return series itself. We extract the “new” part of return for each date and transform the serially correlated return series into one with returns that are uncorrelated over time.

The second approach not only uses the basic series and its serial correlation, it also incorporates useful and relevant information about changes in value from the stock market (or bond market, for distressed debt, or other price series for other portfolios such as oil & gas, etc.). By exploiting the additional information from the relationship between partnership returns and stock market returns, we should be able to get a more accurate picture of portfolio value.

What the first approach adds over the second is that it does not rely on a consistent pattern of value updating. In other words, the regression uncovers the average degree of staleness across time and applies that pattern for updating. If the staleness pattern is not consistent across time, then updating on a company-by-company basis should be able to improve on it. For example, suppose that when the market is rising, companies do a round of funding more often, so the average staleness of pricing is shorter. On the other hand, when the market is falling, the time between deals extends, and prices are on average more stale. Updating on a company basis will capture these differences between rising and falling markets. It will still incorporate useful information from the stock market through the valuation

formulas, which incorporate stock market changes, so it does everything both the first and second approaches do, plus more. It produces a series free of serial correlation (like #3), it uses what's useful from the stock market (like #2) and it uses other information about how much time has passed since each company was valued, whether it was shipping, etc., to update value.

In sum, there are two issues to building a benchmark for private equity assets. First is **whether to build a benchmark that is part stale and hence easily comparable to returns reported to investors** by their general partners, or current. Second is whether the **construction of the benchmark begins with returns on funds or asset valuations**. In building the Sand Hill Index we chose the route of building the index on companies rather than funds, and to seek the goal of an index of all current value instead of part stale values.

The most important factor driving our choice is that it is easier to identify the universe of venture-funded companies than the universe of venture funds. Poorly performing funds are more likely to disappear, leaving no useful data. Second, we believe that it is easier to bring valuations current using company data rather than fund data because we can update company-by-company instead of making some strong assumption about the structure of staleness through time (such as "the entire portfolio is always two quarters stale"). With the company-based approach, staleness can be a function of recent market movements, allowing more staleness as markets descend, less as they rise.

As we see it, the only merit of a part-stale benchmark, such as one that is just averages partnership returns, is that it is easily comparable to the investor's quarterly returns as they are reported by GPs. But as soon as we want to do some statistics to quantify performance more precisely than the eyeballing approach admits, we get into running regressions. As long as we are going to run regressions, we may as well take the small extra effort to do the regressions on a current benchmark, which gives us not only better measures of risk and performance, but the tools for making to market as well.

**No benchmark for alternative assets is going to be investable** the same way the S&P500 is. There are too many funds (a couple of thousand at most) with too few different assets (30 to 50 companies for venture funds) for this to be the case. Nonetheless, we think **this is not a reason to abandon index construction** for the universe of investments. Investors still want to know, and should want to know, how their portfolios are doing compared to the averages in terms of both return and risk. If they are doing better than average, or worse, it would be useful to them to get some insight into why. In addition, in our experience, investors with 40 or 50 partnership interests come in with results that are close to the performance of the Sand Hill Index. This suggests that although the benchmark may not be investable, it is doing a good job of summarizing average performance, and thus it is a useful guide not only to evaluating performance for individual portfolios, but also for generating the necessary statistics for allocating assets and forecasting and managing risk.