A model weighting game in estimating expected returns.

By Lubos Pastor

(published in *Financial Times*, *Mastering Investment*, May 21, 2001, with minor editorial changes)

Lubos Pastor is assistant professor of finance at the University of Chicago Graduate School of Business.

****

It is hard to overstate the importance of expected returns in investments. For money managers, expected returns on various asset classes are key inputs in portfolio decisions. For corporate managers, expected return on their company's stock is a central element of the company's cost of capital, and thereby affects which investment projects the company decides to undertake.

Even ordinary consumers are affected by expected return estimates. The prices charged by many utility companies are regulated to ensure that the utility earns a "fair rate of return", defined by regulatory practice as the utility's cost of capital. How much we pay on our gas and electricity bills therefore partly depends on how the regulators estimate expected returns on utility stocks.

Unfortunately, expected returns are as elusive as they are important. There is no absolute agreement among finance professionals on how expected returns should be estimated. This article compares the relative merits of some common approaches to this challenging task, and argues that the best estimates are produced by combining finance theory with historical returns data and our own judgment.

Does history repeat itself?

"I know of no way of judging the future but by the past."
Patrick Henry

One simple estimator of an asset's expected return is a sample average of the asset's historical returns. Unless we suspect that expected return changes nontrivially over time, the sample average return is an unbiased estimator of expected future return – that is, it is not systematically higher or lower than the true expected value. The unbiasedness of the sample average return is its main advantage.

However, getting things right on average is not the only objective. Walking around the economics departments, you might overhear a joke about three econometricians who went out hunting, and came across a deer. The first one fired a shot, and missed by 10 yards to the left. The second one missed by 10 yards to the right. Instead of firing, the third one shouted in triumph: "We got it!".

The main disadvantage of the sample average is its imprecision. Suppose we want to estimate expected return on the stock of General Motors Corp. (GM), traded on the New York Stock Exchange. Using monthly returns in January 1991 through December 2000, the sample average return on GM is 14% per year. The standard error, the usual statistical measure of imprecision, is huge: 10% per year. With 95% confidence, the true expected return is within two standard errors of the sample average, or between -6% and 34% per year. We want to be more confident than that!

Would the precision increase if we used weekly instead of monthly data? No. Although higher-frequency data helps in estimating variances and covariances of returns, it does not help in estimating expected returns. Intuitively, what matters for expected return is the beginning and ending levels of prices over a given period, but not what happens in between.

The only way to get a more precise sample average is to collect more data further back in time. For example, if we use GM returns all the way back to December 1925, the historical average is 15.5% and the 95% confidence interval narrows to 8.7% to 22.3%. However, the interval is still too wide for comfort.

Moreover, GM today is very different from GM seventy years ago, so the current estimate could be contaminated by old data. In general, as we add older data, we gain precision at the expense of introducing potential bias. Striking the right balance is difficult and requires sound judgment.

Despite its drawbacks, the long-run average return is a popular estimator for expected returns on aggregate market indices. Unfortunately, we have no theory for what the expected market return should be.

Luckily, for individual stocks and most portfolios, we can rely on estimates produced by theoretical asset pricing models. Those estimates tend to be substantially more precise than sample averages.

Theory is good.

Finance theory says that riskier assets must offer higher expected returns, and asset pricing models quantify this insight. The most common pricing model, the Capital Asset Pricing Model (CAPM, pronounced CAP-M), says that a stock is riskier the more closely its price moves with the prices in the market as a whole. The appropriate measure of a stock's risk is therefore the degree of its comovement with the market, which is summarized by a measure called beta (β). Expected return on a given stock in excess of a risk-free rate r_f is then proportional to the stock's beta:

$$E(r) - r_f = \beta \times [E(r_m) - r_f].$$

The constant of proportionality, $E(r_m\text{-}r_f)$, is the expected excess return on the market as a whole, and is often called the equity premium.

What value we choose for the risk-free rate r_f depends on our objectives. If we want to forecast expected stock returns over the next month, the appropriate risk-free rate r_f is the yield to maturity on a Treasury bill that matures in one month. If we want to estimate the firm's cost of capital in order to value the firm's future cash flows, the risk-free rate should be derived from a longer-term Treasury bond, such as a ten-year bond. The bond's duration should come close to the duration of the firm's cash flows. Very long-term bonds should be avoided, though, because their yields might also reflect premiums for risks such as inflation.

Does the above equation solve all of our problems? Not quite. The two quantities on the right-hand side, beta and the equity premium, are not known with certainty -- they need to be estimated.

Beta is typically estimated by regressing the most recent five to ten years worth of monthly stock returns on market returns. For example, the estimate of GM's beta using its monthly data in January 1996 through December 2000 is 1.11.

How much data should we use to estimate beta? The tradeoff is similar to that involved in using sample averages: the further we go back in time, the higher the statistical precision of the estimate, but the bigger the possibility of introducing some old-data bias.

Unlike with sample averages, however, here it often pays to use more frequent data. For example, whereas the 95% confidence interval for GM's beta based on the monthly data is 0.65 to 1.57, this interval based on weekly data is tighter, 0.69 to 1.08. GM's beta estimated using weekly data is 0.88.

However, going from monthly to weekly data is recommended only for the most liquid and volatile stocks. For other stocks, some week-to-week price changes are simply movements between the bid and ask prices around the same true price, which introduces additional error into the estimation. Also, it may take a while for market-wide news to get into the prices of illiquid stocks, which biases the usual beta estimates downward. Conveniently, betas of illiquid stocks can be estimated using an alternative approach developed by the Chicago economists Myron Scholes and Joseph Williams.

It is clear from the GM example that the usual estimates of beta contain a fair amount of noise. A common and useful way of reducing that noise is to "shrink" the usual estimates to a reasonable value, such as one. The value of one is reasonable because the average beta across all stocks is one, by construction.

The "shrinkage" estimate of beta is the weighted average of the usual sample estimate and of the shrinkage target. For example, the "adjusted" betas reported by Merrill Lynch put a 2/3 weight on the sample estimate and a 1/3 weight on the value of one. The adjusted weekly beta for GM is therefore (2/3)*0.88 + (1/3)*1 = 0.92.

Shrinkage betas can be justified as so-called "Bayesian" estimators, in that they reflect not only data but also prior knowledge or judgment. Bayesian estimators have rock-solid axiomatic foundations in statistics and decision theory, unlike many other estimators commonly used by statisticians.

Before seeing any data on GM, we know that it is a stock, so a good prior guess for GM's beta is one. Besides, we typically know more about the company; for example, we know which industry it operates in. Since the average beta among auto-manufacturing companies is around 1.2, a reasonable prior guess for GM's beta is 1.2.

How much weight we put on the prior guess and how much on the estimate from the data depends on the precision of the sample estimate and on the strength of our prior beliefs. Those beliefs can be based for example on the dispersion of the betas of other auto manufacturers: the stronger the concentration of auto company betas around 1.2, the more weight we put on the prior guess. With equal weights on the prior guess and the weekly sample estimate, GM's industry-adjusted beta is (1/2)*0.88 + (1/2)*1.2 = 1.04.

Unfortunately, the CAPM says nothing about expected market return, and estimating the equity premium is more difficult than estimating betas. More frequent data does not help and there is no obvious prior guess. The most common approach is to average a long series of excess market returns, which leads to equity premium estimates anywhere between 5% and 9% per year, depending on the sample period.

A recent equity premium study by Robert Stambaugh from the University of Pennsylvania and the author puts the current premium in the United States at 4.8% per year. This estimate comes from a model in which the premium changes over the last 165 years. Combining this estimate with GM's industry-adjusted beta and a 6% risk-free rate, the CAPM estimates GM's annual expected return (i.e. its cost of equity capital) at 6% + 1.04*4.8% = 11%.

But our models are not perfect…

The CAPM is just a model, not a perfect description of reality. Indeed, many academic studies reject the validity of the CAPM, since some stock return patterns seem inconsistent with the model. Does this mean that we should throw the model away and rely only on model-free estimators, such as the sample average return?

No! Every model is "wrong", almost by definition, because it makes simplifying assumptions about our complex world. But even a model that is not exactly right can be useful.

It is again helpful to adopt a Bayesian perspective and combine what the data tell us with our best prior guess. While the data speak to us about expected return through the sample average return (14% per year for GM), our prior guess can be based on finance theory such as the CAPM (11% per year for GM). The resulting estimate is a weighted average of the two numbers. The weights depend on how strongly we believe in the model and on how strongly the model is violated in the data.

This Bayesian approach is developed in another recent study by Robert Stambaugh and the author. The study finds that even if we have only modest confidence in a pricing model such as the CAPM, our cost of capital estimates should be heavily weighted toward the model. Average stock returns are noisy, so they should receive small weights. In other words, theory is more powerful than data when estimating expected stock returns.

To make the water muddier, the CAPM is not the only theoretical model of expected returns. Serious competition comes from multifactor models, in which expected return depends on the stock's betas with respect to more factors than just the market. The factors can be either macroeconomic variables (e.g. a five-factor model by Chen, Roll, and Ross), or portfolios formed based on firms' characteristics (e.g. a three-factor model of Fama and French), or even return series constructed using statistical techniques such as factor analysis.

Different people have different opinions on which model is the best, and the jury is still out. Meanwhile, what are we to do? A sensible solution is to construct a weighted average of expected return estimates from all models that we are willing to consider, including the "no-theory" model that produces the sample average estimate. Each estimate should be weighted by the probability that its parent model is correct.

Where do we get these probabilities? It helps to be aware of the relevant research, but in the end this is a matter of judgment. The author believes that, despite its weaknesses, the CAPM is the model with the strongest theoretical foundation, and should therefore receive the largest weight. Other models should receive weights commensurate to their theoretical support as well as their empirical success.

Where does the uncertainty come from?

Although pricing models generally produce expected return estimates that are significantly more precise than sample averages, substantial uncertainty remains. Recent research by Eugene Fama from the University of Chicago and Ken French from the Massachussetts Institute of Technology shows that standard errors of more than 3% per year are typical for estimates of industry costs of equity based on commonly used pricing models.

Where does most of the uncertainty come from? Is it more important that we do not know the true beta, the exact value of the equity premium, or that we do not know the right model?

Interestingly, not knowing which model is right turns out to be less important on average than not knowing the parameters within each model. That is one of the conclusions of the author's cost of capital study mentioned earlier. We should therefore spend less time searching for the right model, and more time trying to improve the estimates within each model.

In addition, uncertainty about the premium is bigger than uncertainty about the firm betas, which makes the intangible equity premium the biggest source of uncertainty in the firms' cost of capital estimates.

As popular as they are, asset pricing models are not the only option in estimating the cost of capital. An alternative approach that is often used for regulated utilities in the U.S. is based on the so-called Gordon growth model, wherein the cost of equity equals the sum of the current dividend yield and the long-term dividend growth rate.

This approach is generally favored less by the academics, for several reasons. It makes the strong assumption that dividends will grow forever at the same rate. Besides, there is no theory to help us estimate the dividend growth rate, which is unfortunate because the cost of equity estimate is very sensitive to that rate. This approach therefore strongly reflects subjective opinions about the firm's future prospects.

Shooting at a moving target

There is an emerging consensus in the academia that expected returns vary over time. For example, expected stock returns seem to be related to the business cycle -- they tend to be higher in recessions and lower in expansions.

Among the variables that have been found useful in explaining the time-variation in expected stock market returns are the aggregate dividend-price ratio (D/P) and earnings-price ratio (E/P). Low values of these ratios have historically predicted low returns. In other words, when prices are high relative to the fundamentals, future returns are on average low, especially at longer horizons such as ten years ahead.

The predictive power of the D/P and E/P ratios was reinforced last year, when the S&P 500 lost 10% of its value while the ratios were at their historical lows. However, these predictors worked poorly in the 1990s, when low D/P and E/P peacefully coexisted with high stock market returns.

"If you torture the data long enough, Nature will confess," says Ronald Coase, a Nobel-winning economist. If you search through enough variables, you will certainly find a variable that appears to predict returns. However, this apparent predictability exists by pure chance, and such "data-mined" variables simply will not work in the future.

An interesting example of data mining is provided by David Leinweber, managing director of First Quadrant, a money-management firm in Pasadena, California. Leinweber "sifted through a United Nations CD-ROM and discovered that historically, the single best predictor of the Standard & Poor's 500-stock index was butter production in Bangladesh." (Business Week, June 16, 1997) Good luck if you try to make money on this -- you'll need it. Only predictors with a solid theoretical justification have a chance of working not only in the past but also in the future.

Fortunately, economists have come up with good reasons for why D/P and E/P could have some predictive power, drawing some poison out of the data-mining critique. Some difficult questions remain, though. Is the predictive relation linear? What is the best way to estimate the unknown parameters of this relation? What other predictors should we include?

Of course, if it is hard to estimate expected returns when they are constant, it is even harder to estimate them when they change through time.

There is no simple recipe on how to estimate expected returns. Since data is noisy and no theory is strictly flawless, judgment enters the process at numerous points. There is nothing wrong with that. After all, economic theories themselves ultimately reflect our judgment about how the world behaves.

Given the importance of expected returns and the huge uncertainty associated with them, the finance profession clearly needs to invest more into their estimation. Such an investment will undoubtedly provide a high expected return. But please don't ask me for an exact number.

Further reading:

Fama, Eugene F. and Kenneth R. French, "Industry costs of equity", Journal of Financial Economics, 1997.
Pastor, Lubos and Robert F. Stambaugh, "Costs of equity capital and model mispricing", Journal of Finance, 1999.
Pastor, Lubos and Robert F. Stambaugh, "The equity premium and structural breaks", Journal of Finance, forthcoming.
Scholes, Myron and Joseph Williams, "Estimating betas from nonsynchronous data", Journal of Financial Economics, 1977.