CS-BIGS 5(2: 102-110 http://www.bentley.edu/csbigs


# A new critique of the traditional method for empirically estimating the returns associated with strategies of stock investment: impact of fixing the investment horizon

**Greg Samsa**
*Duke University, USA*

*Much of the empirical literature on behavioral finance involves analyzing historical databases of stock prices. The null hypothesis in these analyses is that risk-adjusted returns are unpredictable, and the alternative hypothesis is that risk-adjusted returns are predictable in a way that is consistent with an insight from behavioral finance. The value investment hypothesis is such an insight, and it predicts that regression toward the mean will cause the returns of poorly performing stocks to be better than average. Empirical tests using historical databases typically involve selecting a portfolio of stocks that satisfy a certain criterion (e.g., low price-earnings ratio) at a given time, holding the stocks for a fixed period such as a year, selling the stocks, calculating the return, and then repeating the process.*

*We performed a simulation of the returns of a mean-reverting stock that might be purchased by a value investor. Mean-reverting stocks tend to have less business risk than average and also exhibit lower volatility in their prices and thus satisfy two operational definitions of having relatively low investment risk. For simplicity, the return on the mean-reverting stock was assumed to be 0, and performing the simulation in the usual way (i.e., with pre-specified dates of purchase and sale) yielded an annual return of 0. However, when the holding time was allowed to vary stochastically -- that is, by holding the stock for a random period of time until a certain absolute level of return is achieved rather than in the usual way – the return was positive, and perhaps would even exceed the returns of stocks with higher risk. Thus, the usual procedure underestimates the returns of stocks whose prices regress toward their means. Moreover this simulation, in the limited technical sense described above, provides a counter-example to the usual relationship between risk and return.*

*The presentation is of particular interest to economists and investors who study stock market returns, and is accessible to readers with an intermediate level of statistical knowledge.*

## 1. Introduction

### 1.0 Working definition of regression toward the mean

The phenomenon of regression toward the mean (RTM) is the basis for many strategies of value investing. Loosely speaking, RTM asserts that stock prices that have fallen below some historical norm are likely to revert toward their historical levels, and that above-average risk-adjusted returns can be obtained from taking advantage of this tendency. Examples of such strategies include buying stocks whose price/earnings ratios are lower than usual, whose price/sales ratios are lower than usual, and whose dividend yields are higher than usual.

The phenomenon of RTM is rather counter-intuitive, and so we begin with an illustration of its cause.

## 1.1 Example of RTM

For clarity of exposition, this example of RTM is taken from a field outside economics. Let $Y_i$ be the results of an intelligence test taken at time "i". Further, assume that on any given day the observed test results are the sum of a true level of intelligence T and a random error term $E_i$. In practice, $Y_i$ is observable, whereas T and $E_i$ are not. Although an over-simplification, much of the "classical measurement theory" used to describe the results of psychometric tests in fact uses this model: $Y_i = T + E_i$.

The next step in developing the classical "errors in variables" model is to account for variation. People vary in their levels of intelligence, and this is accounted for by the distribution of T. For example, T might be assumed to have a Gaussian distribution with mean 100 and standard deviation $\sigma_T=10$. (As will become apparent, the particular shape of the distribution is much less important than its standard deviation.) Under this assumption, approximately 95% of people have an "intelligence quotient" (IQ) that falls between 80 and 120.

The other element of variation in the model is variation within individuals. Recognizing that observed scores will vary from test to test, even among the same person, $E_i$ is assumed to have a Gaussian distribution with mean 0 and standard deviation $\sigma_E=10$. (As will become apparent, the particular shape of the distribution is much less important than its standard deviation.)

The values of $E_i$ are assumed to be uncorrelated with each other, and also uncorrelated with the true value T. For each person, the observed IQ scores should be distributed around the true IQ value of T. For example, for a person with a true IQ value of 120, 95% of observed scores are expected to fall between 100 and 140.

To illustrate the mechanism of RTM, suppose that exactly 1% of all observed values have a measurement error of 20, that exactly 1% of all observed values have a measurement error of -20, and that exactly 98% of all observed values have a measurement error of 0. Moreover, assume that 1,000,000 people have a true IQ of 100, 10,000 people have true IQ values of 120, and that 100 people have true IQ values of 140. The same distributions are assumed to hold for people with below-average intelligence. Table 1 below illustrates what will be observed.

Moving across the rows of the table, each is centered on the appropriate observed value. For example, the row of people with true IQ of 120 is centered on the observed value of 120 – in other words, people with true IQ of 120 are expected, on average, to also have observed IQ values of 120. This is consistent with intuition.

**Table 1.** Simplified illustration of regression toward the mean

| | $Y_i=60$ | $Y_i=80$ | $Y_i=100$ | $Y_i=120$ | $Y_i=140$ |
|---|---|---|---|---|---|
| True IQ=60 | 98 | 1 | 0 | 0 | 0 |
| True IQ=80 | 100 | 9,800 | 100 | 0 | 0 |
| True IQ=100 | 0 | 10,000 | 980,000 | 10,000 | 0 |
| True IQ=120 | 0 | 0 | 100 | 9,800 | 100 |
| True IQ=140 | 0 | 0 | 0 | 1 | 98 |

Now moving down the columns of the table, consider the people with observed test scores of 120. Many more of these people will have true IQ scores of 100 than will have true IQ scores of 140. The same phenomenon works in reverse for people with observed IQ scores of 80 – many more of these people will have true IQ scores of 100 than will have true IQ scores of 60. Moreover, the reason is obvious: so many more people have true IQ values near the mean value of 100 than have true IQ values that are far in the extreme.

Because of the lack of correlation among the various terms in the model, when the intelligence test is re-administered the expected value of $E_2$ is expected to be 0, regardless of the true value and regardless of the value of $E_1$. Accordingly, upon re-administration of the IQ test at time 2, we expect to observe values that are centered on their true values. For people with observed IQ scores of 120, the expected IQ score at time 2 is slightly less than 110 – that is, reflecting that approximately half the individuals in question have true values of 100 and approximately half the individuals in question have true values of 120. This isn't necessarily consistent with intuition – although it isn't surprising that (in general) people with high observed IQ scores at time 1 are smart, what is surprising is that they are not as smart as they look.

This demonstration is not essentially different when utilizing actual Gaussian distributions. Indeed, it is not essentially different when using any distribution that is concentrated towards its center rather than its extremes. The fundamental cause of RTM is the lack of correlations among the various error terms in the model.

## 1.2 Extension to economics and investment

A similarly motivated errors in variables model can be applied to economic phenomena such as corporate performance and stock returns. For example, $Y_i$ might denote annual return on invested capital. Then, the principle of RTM would suggest that those corporations that are earning lower-than-average returns will improve their relative position over time, whereas those corporations that are earning higher-than-average returns will do the opposite. For example, if the average return on invested capital for AMD is 5% and the same average return for Intel is 20%, the principle of RTM suggests that AMD's returns will tend to increase toward 12% whereas Intel's returns will tend to decrease toward 12%. The reason that this phenomenon is counter-intuitive remains the same – as depicted in the example, Intel truly is a better company than AMD – what is counter-intuitive is that it is not as much better as it appears.

Following this same logic, what is expected to happen to companies with lower than average earnings growth and lower than average price-earnings ratios? The answer is that both values will (on average) regress upward toward their population means, thus resulting in superior returns. All value-based strategies utilize this same insight, in one way or another.

## 1.3 Philosophical basis for the errors in variables model

Although simplistic, the errors in variables model is crucial to the development of modern portfolio theory. In particular, if it is assumed that the stock market is a "perfect market" (e.g., with large numbers of well-informed investors, none of whom are large enough to affect prices), then it follows that the current stock price $Y_i$ represents all available information. (Otherwise, arbitrage will push the price toward its true value.) Moreover, it follows that at the next time point stock prices will only change in response to the new information $E_{i+1}$. This new information – for example, a tsunami, a change in government policy -- is unpredictable (i.e., otherwise, it would have already have been discounted by the previous $Y_i$ ). Accordingly, the series of perturbations (i.e., "shocks", "errors") is uncorrelated both with one another and with the true value T. Once this errors in variables structure is assumed, the other main elements of modern portfolio theory (e.g., the capital asset pricing model, the efficient frontier) become logical consequences (Teebagy, 1998).

From a statistical perspective, the result of the errors in variables model is that stock prices follow a "random walk", such as a drunk might take. When comparing the paths taken by the stock prices of various companies, the decisive factor is $\sigma_E$. When $\sigma_E$ is small, as is the case for established companies with dependable prospects, the stochastic path followed by the stock price is merely tipsy. On the other hand, when $\sigma_E$ is large, as is the case for more speculative endeavors, the stochastic path followed by the stock price is so drunk that it is a wonder that it doesn't pass out on the spot.

People prefer the company of social drinkers to that of sloppy drunks, and investors are no exception. Partygoers demand compensation for putting up with drunks – discerning readers can extend this analogy on their own if desired. Investors are no different – they demand higher returns for holding stocks whose variability is higher than average. The investment decision then becomes a trade-off between risk and return – the key question being "How much volatility is an investor willing to accept in order to achieve a desired rate of return?".

In this intellectual structure, volatility in prices (technically, expected volatility in future prices, for which volatility in past prices is often taken as the best available surrogate) is considered to be synonymous with risk. Nothing else about the company matters – not its products, not its management, not its customers – nothing else at all. The rationale is that all this other information has already been accounted for in the current stock price.

## 1.4 Critique

The counter-argument, perhaps best elucidated by Warren Buffet, is in two parts (Buffett, 2001). First, that risk is not a statistical phenomenon pertaining to volatility in stock prices, but instead is a fundamental attribute of the companies in question. For example, the price of the Coca Cola company stock (KO) usually has low volatility, because its product is in high demand regardless of the vagaries of the economic cycle. If the price of KO is suddenly cut from $60 per share to $30 during a stock market panic it will become a bargain. Because they purchased the stock so cheaply, investors will have less "risk" of failing to meet their long-term financial goals (i.e., Buffet's operational definition of risk). However, a believer in modern portfolio theory should conclude the exact opposite. Because prices are assumed to discount all available information, the new price of $30 isn't a bargain – indeed, by definition bargains cannot exist. Moreover, the rapid drop in price has increased the estimate of price volatility – accordingly, the stock has become "riskier" than before and thus should be less attractive, all else being equal.

The second portion of the argument is that savvy investors can identify and profit from temporary mismatches between $Y_i$ and T, such as the drop in price of KO during a market panic. The performance of Mr. Buffet's investments over the years lends his argument additional support.

## 1.5 Design of a critical experiment

On first examination, an empirical test comparing these two theories should be straightforward. One theory holds that stock prices follow a random walk, and thus are essentially unpredictable.

Another theory holds that stock prices are predictable in a certain way. For example, if the theory is that companies with relatively high dividend yields will outperform the market (in a risk-adjusted sense), then the steps in the algorithm to test this hypothesis are as follows:

1. Define the strategy in sufficient detail to be implemented by computer (e.g., select all stocks whose current dividend yield is at least 5%),

2. Using a database, select (and pretend to buy) all stocks with the desired characteristics as of a certain date such as 1/1/2000,

3. Hold the stocks for a specified period such as 1 year,

4. Using the same database, pretend to sell these stocks as of a certain date such as 12/31/2000,

5. Based on the difference between the purchase and sales prices (plus any dividends received), calculate a hypothetical rate of return for the time period in question,

6. Repeat the process for multiple time periods (e.g., 1/1/2001-12/31/2001, 1/1/2002-12/31/2002, etc.),

7. Using the data from all the time periods, calculate an estimated rate of return, plus a standard deviation,

8. Use this latter standard deviation as an input to risk-adjust the estimated rate of return,

9. Compare the risk-adjusted return of the strategy with that of the overall market. If the strategy being tested outperforms the market the theory is supported – otherwise, not.

## 1.6 Critique of step 1 of the critical experiment

Many of the steps in the above algorithm have been criticized. First, defining strategies in sufficient detail to be implemented by computer is not at all simple. For example, suppose that the analyst recognizes that actual value investors don't simply buy every stock with a high dividend yield, but instead try to select those stocks of companies whose dividends are sustainable and growing. (They especially try to exclude companies whose dividends are likely to be cut.) Sustainability might be made operational using a "dividend payout ratio" of dividends divided by distributable earnings. However, any estimate of earnings – including an estimate of distributable earnings – depends on accounting assumptions. If stocks are to be selected by computer, these accounting assumptions cannot be analyzed with as much insight as an actual investor would apply.

Even if earnings could be adequately defined, the problem of earnings volatility remains. When calculating the dividend payout ratio, should the algorithm use the current year's earnings, next year's expected earnings, an average of the last 5 years' earnings, or something else entirely? The upshot of these concerns is that the strategies being tested are gross over-simplifications of how actual investors proceed. Since the strategies that are being tested are pale imitations of those used by actual investors, statistical power will be reduced. In particular, if a study fails to demonstrate the benefit of value investing, does this mean that the strategy of value investing has been refuted or does it only show that what has failed is a poor imitation of value investing?

Apart from the fact that it is such a rich data source, one of the reasons that economists initially decided to study the stock market was to test whether something approximating a perfect market actually exists. Since they hoped that it did, the fact that their tests had low power didn't necessarily concern them.

## 1.7 Critique of steps 2 and 4 of the critical experiment

Actual investors don't necessarily receive the prices provided by databases. For example, for thinly-traded companies there is often a large spread between the bid and asked prices, and those prices often change when an actual trade is attempted. This phenomenon calls into the research that purports to demonstrate the superior performance of small-capitalization stocks. It also calls into question the returns of stocks purchased during much of the Great Depression of the 1930s – since so few shares of stock were traded that the prices quoted in the databases weren't necessarily real. Finally, it calls into question the prices at the height of financial panics – since markets tend to freeze and executing trades isn't necessarily feasible. The common thread of this criticism is that extreme prices in stock market databases aren't necessarily exploitable in real time.

Not all databases are created equal. One common problem with databases is survival bias – for example, suppose that a database is developed in 2010 covering the period from 2000-2009. If the database starts with companies that existed as of 12/31/2009, it will have deleted all of the companies that went bankrupt during the decade in question. The performance of investment strategies that buy stock in companies at risk of bankruptcy will be artificially inflated.

## 1.8 Critique of step 5 of the critical experiment

Actual investment returns involve slippage, some sources of which include commissions and taxes. These are particularly problematic for strategies that involve lots of short-term trades.

## 1.9 Critique of step 6 of the critical experiment

Although conceptually straightforward, selecting the years to be compared can be problematic. From a statistical perspective, the larger the sample size, the better the resulting inference is. On the other hand, the earlier the period in question, the greater the difference is between the data being used and the present-day market. "How far back the analysis should go" is a matter of debate.

## 1.10 Critique of step 7 of the critical experiment

A technical issue is whether, once results are obtained for each year in question, they should simply be averaged (i.e., arithmetic mean) or whether a geometric mean should be used instead. The usual (i.e., arithmetic) mean is the easier to calculate, but the geometric mean more closely approximates the returns that investors will actually receive.

## 1.11 Critique of step 8 of the critical experiment

The question of how to properly risk-adjust the returns obtained by the above algorithm is a matter of debate among economists, and is not considered in detail here. As an example of the issues being discussed, suppose that an "ugly" strategy of selecting small companies at significant risk of bankruptcy appears to provide superior risk-adjusted returns. Most investors would shy away from implementing such a strategy, if for no other reason than it would be difficult to exit the trade in case they unexpectedly needed their money. Such a "liquidity risk" might not be adequately captured by the risk-adjustment procedure, thus over-estimating its performance.

## 1.12 Critique of step 9 of the critical experiment

Although more of a technical issue than an insurmountable methodological problem, there still remains the question of determining the proper standard of comparison (e.g., large-capitalization stocks, mid-capitalization stocks, American companies, international companies) for the strategy being tested. Ideally, this comparison should also reflect the elements of slippage noted above.

As described, the above algorithm truly is a critical experiment in the sense used by Popper and other philosophers of science – the hypothesis that the strategy in question outperforms the market can potentially be falsified by the data. In other words, if the results of step 9 are negative, the experiment stops. However, in practice if one version of the investment strategy fails, the researcher is likely to tinker with it and re-test. Often, many different versions of the strategy are tested and, indeed, some analysts assert that this is a virtue and that the final strategy is thus "optimized". However, it is well-known within the statistical community that this approach consistently leads to "over-fitting", and that the performance of such "optimized" models will be grossly over-stated. Not only are "past results no guarantee of future performance", past results are almost guaranteed to over-state future performance. Combining over-fitting with leverage led to the Long Term Capital Management collapse, among other financial disasters.

## 1.13 A new critique

The discerning reader will have noticed that the literature contains criticisms of every step in the above algorithm with the exception of the innocent-sounding step 3 (hold the stocks for a specified period such as a year). Apart from being intuitively natural, the original motivation for fixing the holding period was based on economic theory. Specifically, stocks are often held until a specific date – for example, an "investment horizon" might close on the date of a scheduled retirement, on the date that the investor expects to enter college, etc. Much of investment theory begins by fixing the investment horizon.

However, in actual practice investors seldom buy and sell stocks on pre-specified dates. They try to time their buying in order to obtain bargains, and they try to wait until the stocks they hold become over-valued before they sell. (If they need to withdraw money at the same time as a stock market crash, these investors might borrow money using their stocks as collateral rather than selling them at poor prices.) If they succeed in "timing the market", their returns will exceed those of the strategy of buying and selling at pre-specified dates. Indeed, the greater the volatility in the stocks that they hold, the greater the potential increase is in returns associated with buying and selling opportunistically, and the greater the loss is in statistical power associated with performing assessments in the usual fashion.

## 2. Methods

### 2.1 Safe haven companies

The debate between supporters of modern portfolio theory and those of value investing basically amounts to one between nihilists and fundamentalists. The assertion that all available information is contained in the current stock price is essentially nihilistic – as it asserts that the fundamental elements of the company (e.g., its business model, its management) don't matter. Fundamentalists argue the reverse – not only does this information matter but, although it might be ignored by the market in the short-term, truth about long-term corporate performance will eventually win out.

There is one point about which these two schools of thought agree – namely, that stocks of "safe-haven" companies such as Coca Cola (KO), Johnson & Johnson (JNJ) and Consolidated Edison (ED) have lower risk than average. The nihilists base their opinion on the low level of volatility in their stock prices. Fundamentalists base their opinion on the fact that these are large, well-financed companies that offer products that consumers must continually repurchase (e.g., soft drinks, electric power).

In practice, the prices of safe-haven companies evidence a stronger tendency to regress towards the mean than do most others. One explanation for this phenomenon is that, because their stream of earnings and dividends is so predictable, their "intrinsic value" can be estimated with a much higher degree of accuracy than is the case for other companies. Indeed, the prices of these stocks most often depart from this intrinsic value because of the view that large institutional investors have about the other stocks in the market. When these investors are feeling speculative, they sell the safe haven stocks in order to raise money to buy other companies. When these investors are stricken with panic, they do the opposite. At any point in time, there are always at least a few investors that serve to push the prices of these stocks back toward their equilibrium value. Most particularly, when the stock market engages in one of its bouts of speculative excess, there are always some investors that are looking to pick up stocks like KO, JNJ and ED on the cheap. For example, if it is generally agreed that the intrinsic value of KO is $50 per share, these investors would become somewhat interested if the price drops to $45, more interested if the price drops to $40, extremely interested if the price drops to $35, etc. This amounts to a tendency for historically low prices to regress toward their mean – a trend that becomes increasingly strong as the price diverges from its intrinsic value.

The goal of "safe-haven value investors" is to buy safe-haven companies at prices below their intrinsic value, to hold these companies until their prices exceed their intrinsic values, and thus to obtain superior risk-adjusted returns. Nihilists deny that it is possible for any strategy to obtain superior risk-adjusted returns. Empirical testing, using the above-described algorithm, hasn't been definitive. In general, and consistent with a broader literature on behavioral finance, such value-based strategies do tend to outperform the market, but by only a modest amount. This relative outperformance supports the positions of both parties. The fundamentalists point to the outperformance and, indeed, an entire industry has developed around the idea of behavioral-financed-based investment. The nihilists point to the modest levels of outperformance as evidence of how difficult it is to "beat the market", and hold out the possibility that more sophisticated methods of risk-adjustment would cause the apparent outperformance to disappear altogether.

### 2.2 Mathematical model for safe-haven stocks

As a model of the behavior of the prices of safe-haven stocks, assume that the price at time "i" depends on three factors: (a) the previous price $Y_{i-1}$ ; (b) a random error term $E_i$ (uncorrelated with either T or the random errors at other time points); and (c) a term $\alpha$ (with $\alpha > 0$) representing the impact of regression toward the mean. Specifically:

$$Y_i = Y_{i-1} + \alpha(T-Y_{i-1}) + E_i.$$

When the price at time i-1 is less than intrinsic value, $(T-Y_{i-1})$ will exceed 0, as will the term $\alpha(T-Y_{i-1})$. The impact of $\alpha(T-Y_{i-1})$ will be to push the next price toward intrinsic value.

The role of $E_i$ is consistent with that of the errors in variables model – it represents the impact of unpredictable events on the stock price. Indeed, the most important difference between the current model and the errors in variables model is the tendency of prices to regress toward their mean.

Value investors, in general, attempt to purchase stocks when $(T-Y_{i-1}) > 0$, and it is straightforward to demonstrate that doing so would achieve superior returns. However, consistent with the nihilists, we assume that it is impossible to consistently identify when $(T-Y_{i-1})$ exceeds 0 and, instead, assume that the investor purchases stocks when their price equals their intrinsic value $(T=Y_{i-1})$.

Accordingly, we have assumed that investors have no particular edge in stock selection, only that they can accomplish the straightforward task of identifying mean-regressing safe-haven stocks, and that their decisions about which mean-regressing stocks to select are no worse than average.

Finally, we assume that the supply of safe-haven stocks selling at no more than their intrinsic value is unlimited. (The plausibility of this assumption is discussed later.) In particular, we assume that if investors decide to sell a safe-haven stock that they own, there is always another one that can be bought.

### 2.3 A strategy for trading safe-haven stocks

Assume that the investor has a relatively long investment horizon such as 10 years, each of which consists of 250 trading days. On trading day 1, the investor purchases the stock at its intrinsic value. The holding period is indefinite – the investor will sell the stock when its return is R.

For example, if R=4% and the stock is purchased at $100, the investor will sell the first time the stock price reaches $104. Once the stock is sold, another stock with similar characteristics is bought and the process is repeated. At the end of the investment horizon, the stock that the investor owns is sold, regardless of price.

### 2.4 Expected returns from the strategy

The intrinsic value of actual safe-haven companies is expected to increase over time, due to the compounded growth of their earnings. Here, for simplicity, we have assumed that the intrinsic value remains constant.

Accordingly, if the nihilists are correct the expected return of this strategy, estimated by following the returns of a hypothetical large cohort of investors each having an investment horizon of 10 years, should be 0.

On the other hand, if the observed return exceeds 0, we will have demonstrated that the returns of low-risk companies can be enhanced by taking advantage of the stochastic nature of their prices. Primarily, this will serve to illustrate the new critique of step 3 of the testing algorithm. But secondarily, and in the limited sense defined above, this will also provide a counter-example to the investment maxim that in order to obtain increased gains investors must always accept increased risk.

## 3. Results

Table 2 presents results under the following scenarios. The investment is purchased at the intrinsic value of $100. The distribution of the errors is Gaussian with mean 0 and standard deviation 1. The parameter representing regression toward the mean, $\alpha$, is 0.01. The investment horizon is 10 years (2,500 trading days). If the price exceeds the threshold value, a profit is taken and the simulation is re-set – that is, another stock is purchased at $100. On day 2,500 the stock is sold. For example, setting the threshold value to $101, on average the value of the investment at day 2,500 is $224, for an absolute return of $124 (i.e., $224 minus the original $100).

For the above set of parameters, as the threshold value decreases the absolute return increases. For a threshold value of $101, the absolute return of $124 corresponds to an annualized return exceeding 8%. Please note that this $124 is an excess return – that is, an excess return above the 0% that would be expected from a buy-hold strategy, since it is assumed that the intrinsic value is unchanged over the course of the simulation.

**Table 2.** 10-year absolute returns, keeping the purchase price of $100 constant, varying the threshold value

| Threshold value | Absolute 10-year return |
|---|---|
| $101 | $124 |
| $102 | $117 |
| $103 | $109 |
| $104 | $101 |
| $105 | $93 |

Table 3 illustrates that the impact of the “noise-harvesting” element of this strategy exceeds that of a “stock-selection” element. Purchasing mean-regressing stocks with intrinsic values of $100 for $95 leads to an absolute return of $163, but purchasing the same stocks for $100 still leads to an absolute return of $124.

**Table 3.** 10-year absolute returns, keeping the threshold value of $101 constant, varying the purchase price

| Purchase price | Absolute 10-year return |
|---|---|
| $99 | $132 |
| $98 | $140 |
| $97 | $147 |
| $96 | $154 |
| $95 | $163 |

## 4. Discussion

Our primary purpose here is to illustrate a new critique of much of the literature on investment returns – namely, that by assuming that investors hold stocks for a fixed period of time, their returns will be underestimated. This critique only holds for the subset of the stocks in the market that exhibit mean reversion. Nevertheless, and only in a limited technical sense, this phenomenon provides a counter-example to the usual relationship between risk and return. Mean-reverting stocks have less business risk than average, and also exhibit lower volatility in their prices. Nevertheless, "harvesting their volatility" can increase their returns notably.

Our demonstration has a number of limitations. First, the model for the price behavior of mean-regressing stocks is simplistic. For example, it does not address the phenomenon of price momentum. In particular, momentum would serve to undermine the process of regression toward the mean during the period for which momentum causes a stock's price to increasingly diverge from its intrinsic value.

Moreover, our distributional assumptions do not reflect the possibility of occasional price crashes. For example, the oil giant BP would have been a natural candidate for mean reversion, right up until the Gulf oil spill. This is particularly problematic as the strategy being tested is "short volatility" – and being "short of" anything carries within it a non-trivial set of dangers.

Second, the parameters of our model were chosen for purposes of illustration, and were not empirically estimated. In particular, if the variability of the error term E in our model is unrealistically large, so will be the impact of regression toward the mean. Our simulations demonstrate the presence of an effect, but do not necessarily estimate its magnitude.

Third, since this paper is not primarily concerned with the empirical performance of actual investment strategies, we have not engaged in the traditional data-based assessment of the returns associated with the above strategy. Interested readers are encouraged to perform such a test, if desired.

Fourth, embedded within our model is a strong assumption – namely, that an identically-performing replacement can be found whenever a stock is sold. In one sense, such an assumption is unrealistic as the most extreme mean-reverting stocks (e.g., the "refrigerator and medicine cabinet stocks") evidence a strong correlation among their prices. On the other hand, such an assumption is consistent with the well-documented phenomenon of "sector rotation" – presumably, an investor could always find an out-of-favor sector and then select a high-quality company within that sector under the assumptions that: (a) the current stock price is probably no higher than its intrinsic value; and (b) because the company is of high quality, it will eventually be in sufficient demand among investors to cause the phenomenon of regression toward the mean.

Finally, when taken literally, the strategy being tested would suffer considerable slippage due to commissions and taxes. However, it is equally applicable to longer time frames. Presumably, the investor would want the time frame to be long enough to comfortably avoid over-trading, yet short enough so that there remains volatility to be harvested. A medium-term implementation of this strategy might involve writing covered call options. Such options would: (a) immediately monetize the impact of price volatility; (b) minimize the impact of being short volatility by buying only high-quality companies whose prices are expected to recover from temporary declines; and (c) not worry the investor when the stock in question is called away, as this imply that a large (and presumed temporary) rise in the stock's price (e.g., above the stock's intrinsic value) has already taken place.

In summary, investors (as differentiated from pure speculators) all attempt to buy stocks at reasonable prices and must all cope with price volatility. Value investors hope that volatility will temporarily push the price of stocks so low that they can buy with a margin of safety (i.e. *at far below their intrinsic values*). Whether this can be done consistently is a subject of debate. What we have contributed to this debate is a counter-example – demonstrating that excess risk-adjusted returns are possible by using a volatility-harvesting strategy *even when stocks are purchased at their intrinsic values* (so long as their prices exhibit the tendency to regress toward their means). Of course, nothing in this paper constitutes actual investment advice, and the fact that we have demonstrated this counter-example (albeit under restrictive assumptions) doesn't necessarily imply that readers should try this with their own money.

## 5. Comment

If, for the sake of argument, the core message of this manuscript is accepted, at least two questions naturally follow. The first question pertains to the sociology of modeling: "why hasn't it been noticed that one of the standard models of investment performance imbeds an unrealistic assumption?". Discussed in detail elsewhere (Samsa, 2013), the short answer is that (a) the assumption of a fixed time horizon was derived from another context; (b) it is usually satisfactory, mean-reverting stocks being a special case; and (c) economists don't usually trade stocks for a living, and thus won't have direct experience with the subtle nuances of the system being modelled.

The second question is "when would an investment strategy based on the principle of RTM be most likely to fail?". The answer to this question, anticipated by the previous BP example, is "when the intrinsic value of the stock being purchased drops significantly and permanently". Given the dynamic nature of the economy, an implication of which is that "safe havens" are only relatively so, this suggests that a RTM-based strategy would be most likely to fail within a long- rather than a short-to-medium timeframe. This makes Warren Buffett's performance in buying safe-haven companies and holding them for the long term all the more remarkable.

Correspondence: samsa001@mc.duke.edu

## References

Buffett, W.E. (au), Cunningham LA (ed). The Essays of Warren Buffett: Lessons for Corporate America. Cunningham Group, 2001. ISBN-13 9780966446111

Nunnally, J.C., Bernstein, I.H. Psychometric Theory. McGraw-Hill, 1994, ISBN-13 9780070478497.

Samsa, G.P. An example of an investment model that makes something out of nothing (sort of): implications for building applied mathematical models. Applied Mathematics. 2013:4;89-96.

Teebagy N. *The Math Behind Wall Street: How the Market Works and How to Make it Work for You*, Four Walls Eight Windows, New York NY, 1998.