Market truths: theory versus empirical simulations

WILLIAM C. WOJCIECHOWSKI* and JAMES R. THOMPSON
Rice University, 6100 Main Street, Houston, TX 77005, USA

(Received 1 August 2004; in final form 3 October 2004)

There have been many debates over the true nature of financial markets and the implications regarding passive and active portfolio management. Much of the literature employ advanced quantitative techniques to produce evidence against, or for, a developed theory. Often times, these studies make strict assumptions and judgment calls that allow one to question the validity of the results. In this paper, we combine computer simulation and exploratory data analysis to gain a better understanding of stock market portfolios and of common risk measures. By taking a computational approach, our study requires a small amount of subjectivity and makes no assumptions about capital markets. This makes our results very difficult to discredit. Our results provide strong evidence against the efficient market hypothesis and display some undesirable properties of several well-known risk measures.

Keywords: Simulation; Exploratory data analysis; Portfolio selection; Market portfolio; Efficient markets; Risk measures

1. Introduction

Over recent years, index fund investing has shifted large amounts of funds from the active portfolio manager to the hands of the passive manager, see ref. [1]. Is it the intrigue of a simplistic strategy, distrust of an industry bred by the criminal acts of a few, the herd mentality, or a combination of these, and possibly other factors, which leads financial product consumers to index funds? Furthermore, the implications of the seminal works by Markowitz [2], Fama [3–5], Sharpe [6], Lintner [7], and Mossini [8] are seductively simple to understand and offer market behavior theories that support passive investing. In all of these theoretical explanations, an important detail has been overlooked. Do empirical observations of the stock market agree with the theory? When these theories were developed, there were not enough readily available computational resources to observe stock market portfolios using computer simulation. Because of this, one must ask, would different theories have been developed if a computational approach could have been used?

The developed theory supports passive investing, and our study determines whether active investing has the possibility of outperforming passive investing. Furthermore, not only does
passive investing involves passive allocation, but it also implicitly employs passive risk management. The end result is a portfolio with no risk management. This detail is truly the Achilles’ heel of market funds, and actively managed funds would do well to exploit this fact when promoting their business. The recent bear market is timely in that it shows the magnitude of losses that can be incurred without risk management. Having noticed that index funds have no explicit controls for risk, we were curious to see whether the market portfolio truly offered optimal risk adjusted return. It would indeed be strange if a portfolio lacking risk controls is as good as it gets. In this paper, we provide simulated empirical observations of portfolio behavior that challenge common notions regarding the relative performance and risks of stock portfolios that are actively and passively managed.

Because the performance of a portfolio must be adjusted for risk, we also did simulations to study the behavior of several common risk measures. We wanted to see whether these common measures could be used to make a rational ordering between risky and less risky instruments. Although the results are not technically surprising to the quantitatively trained, the results are an example of how basic simulation results can show non-quantitative people the deficiencies of several common, and widely employed, risk measures.

2. Portfolio simulation methodology

Our research is different in that we do not directly address the issue of efficiency by testing for correlations in the fashion of Fama [3], Lo and MacKinlay [9], and Hagerman and Richmond [10]. Because there is a large body of the literature arguing for both sides, as is indicated by the texts of Lo and Mackinlay [11] and Malkiel [12], we do not pursue this type of research. Furthermore, such tests are reliant upon time horizons, market segment, and time period, adding a taint of subjectivity to the procedures. This is easily seen by the abundance of works that support and refute the existence of correlations in capital markets. Although we select portfolios and measure their performance, it is not our point to support our results by examining specific trading strategies. There are many works that address this subject and we direct the reader to Pruitt and White [13], Ball et al. [14], or Bessembinder and Chan [15] for different views on this subject. Because of the difficulty of developing and testing trading strategies, we chose to avoid this route. In addition, even if one finds a profitable trade, it is not necessarily true that the strategy is appropriate for every investor. Our work is similar to that of Haugen [16] in that we find that the market portfolio is not optimal and that volatility and returns are not positively correlated. This work is much different in that we do not make strict statistical assumptions and judgment calls found in such works. Some weaknesses of such classical approaches include assuming normality of returns, removing outliers greater than four standard deviations from the mean, observing the performance of only one active management style and employing a method such as ordinary least squares (OLS) that is extremely sensitive to outliers and departures from assumptions. All of these deficiencies can bias the results and possibly generate false conclusions. Given the pitfalls associated with the techniques described previously, we decided that a novel approach would better suit our purposes and provide results that are more robust and more difficult to refute.

What we are really interested in are the possibilities for the active portfolio manager. For instance, if we found that no portfolios outperform the market on a risk adjusted basis, then we can expect that an active manager will under-perform the market, and the incurred fees are payment for tailoring a portfolio to the client’s needs. Therefore, we argue that the true measure of a portfolio manager’s performance is not relative to the market but how well their execution matches the stated plans. The stated plans should quantify performance in both terms
of performance and terms of risk measures. Here, we obviously assume that the investment plan is sound. Surely, the manager’s remuneration should be directly tied to how well the manager achieves the client’s goals. This arrangement aligns the manager’s incentives with her fiduciary responsibilities. All of this sounds very obvious, but it has other implications for judging the manager’s performance. On that regards, our methodology also allows one to examine the feasibility of a manager’s performance claims or to determine whether a client’s goals are obtainable.

Our approach is to look at a multitude of scenarios and to see what type of performance is actually possible. By computer simulation, we examine thousands of portfolios and determine what type of stock portfolio performance is obtainable and compare the performance of these portfolios to the market portfolio. It is important to note that we are not simulating a trading strategy. Also, we would like to emphasize that we are not advocating that random portfolio selection is a valid investing plan. Instead, we randomly generate many portfolios and observe the empirical market performance of each portfolio. Because each of the random portfolios is one that could possibly be held, we observe what type of performance could have been obtained.

Because the method is simplistic, it is easy to overlook the power of such an approach. An exercise such as this gives us the knowledge to state several things. First, without any trading rule specification, we are able to determine whether there should be any trading rules, or active management styles, that will outperform the market. Also, at the same time, we can observe the amount of out or under-performance that active management could possibly achieve relative to the market. It is also important, and possible with our methodology, to observe what types of risks were incurred to achieve such performance. This third point is a complete study in itself. Because of this, we concentrate on the first two items. These are not the only areas that this type of approach can illuminate. However, it is possible to generate an overwhelming amount of information, and we chose to focus on a few points.

Let us begin by describing our procedure for simulating portfolios. The time period we chose was from 1970 to 2002. This period was selected to allow a wide spectrum of market conditions to enter the study. Also, because of data availability, this time period was chosen to allow a large enough universe from which to select random portfolios. Each year, the universe consisted of the largest 1000 market cap companies as of January, 1 of that year. From each universe, we generate 50,000 random portfolios. The random portfolio generation process is as follows:

1. Randomly generate \( w_j^* \) for each company from a standard uniform distribution, \( j = 1, \ldots, 1000 \).
2. Let \( w_j = w_j^* / \sum_{j=1}^{1000} w_j^* \).

Using this process, \( 0 \leq w_j \leq 1 \) and \( \sum_{j=1}^{1000} w_j = 1 \). After generating the random weights, we calculate the performance measures: the annualized return and annualized volatility of the random portfolio. This calculation proceeds as follows:

1. Let \( r_{j,i} \) be the return of company \( j \) on day \( i \).
2. Let \( w_j \) be the random weight of company \( j \), generated from the process described previously.
3. Let \( N \) be the number of trading days in the year.
4. Let \( P_{s,i} = \sum_{j=1}^{1000} w_j r_{j,i} \) be the return of the random stock portfolio on day \( i \).
5. Let \( \mu_s^* = \sum_{i=1}^{N} P_{s,i} / N \) be the average daily return of the random stock portfolio.
6. Let \( \mu_s = \mu_s^* N \) be the annualized random stock portfolio return.
7. Let \( \sigma_s = \sqrt{[N/(N - 1)] \sum_{i=1}^{N} (P_{s,i} - \mu_s^*)^2} \) be the annualized volatility of the random stock portfolio.
To determine whether or not the random portfolio has under or outperformed the market, we compare the point \((\sigma_s, \mu_s)\) to the Capital Market Line. As a brief review, the Capital Market Line contains all volatility and return pairs that correspond to a portfolio constructed from a linear combination of the market portfolio and a riskless instrument. Specifically, the return from a portfolio on the Capital Market Line is constructed as 

\[ P_c = \alpha P_m + (1 - \alpha) \rho, \]

where \(P_m\) is the market portfolio return, \(\rho\) is the riskless instrument return, and \(\alpha > 0\) is a constant representing the allocation between the market portfolio and the riskless instrument. Note, \(\alpha > 1\) is interpreted as borrowing, or leveraging, to invest in the market portfolio. We used all stocks on the NYSE, NASDAQ, and AMEX, weighted by market cap, as the market portfolio and the 1-year treasury bill as the riskless instrument. Letting \(\mu_m\) be the annualized market portfolio return, \(\sigma_m\) be the annualized market portfolio volatility, \(\rho\) be the interest rate of the 1-year treasury bill at January of each year, and knowing that the points \((0, \rho)\) and \((\sigma_m, \mu_m)\) lie on the Capital Market Line, we obtain a Capital Market Line by finding the line that passes through \((0, \rho)\) and \((\sigma_m, \mu_m)\). The line equation is 

\[ \mu = \rho + (\mu_m - \rho)/\sigma_m \sigma. \]

Therefore, all portfolios with a return of \(\mu\) and a volatility of \(\sigma\), satisfying this equation, are on the Capital Market Line. If a random portfolio falls above the Capital Market Line, we say the random portfolio has outperformed. Otherwise, it has underperformed. We measure the amount of out performance by the distance a point is above the Capital Market Line.

3. Simulated portfolio results

We now present evidence that it is incorrect to invest in the market portfolio because it is optimal. Here, we define optimal as selecting a portfolio on the Capital Market Line. Observing figure 1, the percentage of portfolios that are above the Capital Market Line ranges from close to 0.0% to >90.0%. Also, in 66% of the years, ≥50% of the random portfolios are above the Capital Market Line. To see that 66%, or 22 out of 33 years, is larger than expected by random chance, let \(X\) be a binomial random variable with \(n = 33\) and \(p = 0.5\). Then, 

\[ P(X \geq 22) = 0.04. \]

This is hardly what we would expect if passive investing were truly optimal. If random selection can exhibit such behavior, then a competent, professional portfolio manager should be able to do as well as, if not better than the random selection. Again, we would like to make it clear that we are not promoting random selection as a valid investment strategy. The proclamation here is that if random selection can do better than a market cap weighted

![Figure 1](image-url)
portfolio, then a competent active portfolio manager should also be able to outperform the market cap portfolio. Also, this manager should be able to offer a layer of risk management that is not possible with a market portfolio. All in all, the evidence suggests that reasoning by optimality does not justify passive investing, especially if one considers that passive investing offers no risk controls.

It is also of interest to see what type of market conditions allows the portfolio manager to benefit the investor the most. To do this, we examined the average out-performance for both bear and bull market years. When the market is up, the median out-performance is 2.5%, the lower quartile is 1.8%, and the upper quartile is 3.9%. Now, this is from selecting randomly. An effective portfolio manager could possibly provide better results. The bull markets look promising, but even more amazing are the bear market results. In this case, the median out-performance is 5.2%, the lower quartile is 3.1%, and the upper quartile is a staggering 8.7%. Therefore, the portfolio manager can add much value by managing the downside during the bear years. Or put another way, much value can be added through risk management. Therefore, a successful portfolio manager should be adept at measuring, monitoring, and managing risk. Furthermore, the manager’s remuneration should be based not only on his or her ability to generate profits but also on risk management skills. These results provide superiors and investors alike with information to gauge the portfolio manager’s performance. If performance is measured by returns alone, the bear years provide the necessary evidence to identify superior portfolio managers. Obviously, the returns should be adjusted for risk when measuring performance. However, the risk has to be defined and each investor’s perceptions of, and tolerances for, risk will be different. For instance, one client may feel that large down months are uncomfortable, even if the portfolio does well on an annual basis. Another client may wish to avoid exposure to certain sectors. Clearly, there is a multitude of possible risk sources, and an advantage is that active management can eliminate, or at least alleviate, the exposure to such sources. In contrast, when using passive management, the investor gets the exposures that the market gives them. We now turn to the other commonly held belief related to passive investing, the relationship between returns and volatility.

Having shown that there are possibilities for active investing to offer superior performance, we now address another commonly held but incorrect notion about investing. It is ubiquitous to state that greater return equates with greater risk. To hold with convention, we measure risk as volatility. To avoid any confusion, let us state again, we do not believe volatility truly measures risk. We were curious whether the convention held, and we tested this by plotting the annual return and annualized volatility of the random portfolios for each year. Included subsequently are example graphs from the years 1999, 1996, and 1984. Turning to figure 2, we see that, as commonly believed, more return generally implies greater volatility. Drawing attention to figure 3, it is seen that return and risk are not related. Finally, figure 4 demonstrates that more volatility implies less return. Unfortunately, trying to garnish more return by incurring more risk, as measured by volatility, can lead to undesirable results. One might object, stating that 1996 and 1984 were anomalies. To refute such a claim, we present the following statistics. Out of 33 years, 10 have more volatility which implies less return, 13 have no relationship between volatility and risk, and 10 have more volatility which implies more return. Such evidence demonstrates that more volatility cannot be equated with more return. One might even be as bold to state that any investing strategy built around such a notion is fundamentally flawed. This should not be confused with the fact that, to have profit-making opportunities, capital markets must experience some volatility. Most clients would prefer to have the possibility of a larger return on investment, if more risk is incurred. Clearly, this is not the case when measuring risk with volatility. This leaves another matter open for investigation: which risk measures, if any, are positively correlated with returns?
390

W. C. Wojciechowski and J. R. Thompson

Figure 2. Plotted above are the annual returns and annualized volatility of each of the random portfolios generated for the year 1999. For this year, it is seen that more volatility implied greater returns.

4. Review of risk measures

There are many types of risk, market risk, credit risk, and liquidity risk, to name a few. Although it is easy to identify something as being a risk, it is not always clear how to define risk. This is true even though risk has been acknowledged and pondered over the ages. The reader is invited to read Bernstein [17] for an enlightening and informative history of risk. It would seem that a concept that is difficult to define would be even more elusive to measure. This is true as the existence of the multitude of risk measures supports this claim. Nonetheless, for this article, we will roughly define risk as being the chance that an investment’s return will be less than expected sometime in the future.

However, the details of formulating a risk measure are not so clear. For instance, if $X$ is a random variable representing the value of an investment at some future time $t$, then how can $X$ be used to quantify risk? Should $X$ be a scalar such as the standard deviation (volatility) of $X$? Would a vector such as a value and probability associated with that value be better? Or, should one consider the whole distribution of $X$? In this article, we perform simulations to study the behavior of several common risk measures. We chose a subset of the risk measures

Figure 3. Plotted above are the annual returns and annualized volatility of each of the random portfolios generated for the year 1996. For this year, no relationship between volatility and returns is present.
Market truths: theory versus empirical simulations

Figure 4. Plotted above are the annual returns and annualized volatility of each of the random portfolios generated for the year 1984. For this year, it is easily seen that incurring more volatility implies less return.

from Guy and Kaplanski [2002]. The set was chosen to demonstrate how certain classes of risk measures behave.

To begin with, we have the class that measures risk by quantifying the dispersion of the random variable $X$. The two we chose to look at were the volatility

$$
\sigma_x = \sqrt{\int_{-\infty}^{\infty} (x - \mu_x)^2 f(x) \, dx}
$$

and the coefficient of variation

$$
cv = \frac{\sigma_x}{\mu_x},
$$

where $\sigma_x$ and $\mu_x$ are, respectively, the volatility and the expected value of $X$. These measures are well-known and commonly employed. However, there are two flaws with these measures. First, as seen previously, volatility is not positively correlated with returns. This implies that taking on more risk by increasing volatility does not necessarily generate more return. The second flaw is that these measures will penalize investments whose $X$ has more volatility, which is due to a positive skewness of $X$. In plain words, it penalizes large gains as much as large losses.

Seeing the deficiencies of dispersion measures, the next natural risk measure to consider are measures of dispersion but only considering the dispersion below a value. These measures include semi-variance, worst case scenario, and simply the probability of being below a selected value. All of these measures belong to the Fishburn family that is indexed by $\alpha$. The Fishburn family satisfies

$$
\int_{-\infty}^{q} (q - x)^\alpha f(x) \, dx,
$$

where $q$ is the selected value and $\alpha$ is a parameter that selects a specific member. When $\alpha = 0$, the measure is merely the probability of being below $q$. The problem here is that this ignores the size of the loss. When $\alpha = 2$, the measure is the well-known semi-variance. Although the size of loss is accounted for, the probability of being below $q$ changes for each risk instrument we want to consider. The worst case scenario is selected by setting $\alpha = \infty$. Note, this is simply a measure of maximum loss. Because it ignores the probability of such a loss, it is the opposite extreme of $\alpha = 0$. 
No study of risk would be complete without including the famous, and widely used by investment banks, Value at Risk (VaR). VaR is defined as

\[ q - X(P), \]

where \( q \) is a selected value such as zero, or the \( E[X] \). Here, \( X(P) \) is the \( P \)th percentile of the distribution of \( X \). The problem here is that VaR ignores the values below \( X(P) \). Also, consider the case where one is ordering investments from least to most risky using VaR. If the value of \( P \) is changed, the ordering can also change. This is clearly an undesirable property. To alleviate the sensitivity to the choice of \( P \), the accumulate VaR was introduced. It is defined as

\[ \text{AVaR}_q = \int_0^P \text{VaR}_q(z) \, dz. \]

This is not a complete solution as the simulation results will show.

5. Risk measure simulation methodology

Our study considers two investments whose return values follow different distributions. Distributions are chosen so that there is a clear distinction between which investment is more risky than the other. We consider three scenarios: the possibility of a large gain, moderate chance of a small loss, and small chance of a catastrophic loss. For the possibility of a large gain, we use the following normal-mixture densities to generate the investment returns

\[ f(x) = \phi(x; \mu, \sigma) \]
\[ f(y) = 0.95\phi(y; \mu, \sigma) + 0.05\phi(y; \mu + 6\sigma, \sigma). \]

Here, \( y \) is the less risky investment. The moderate chance of a large loss uses

\[ f(x) = \phi(x; \mu, \sigma) \]
\[ f(y) = 0.95\phi(y; \mu, \sigma) + 0.05\phi(y; \mu - 6\sigma, \sigma). \]

The better investment here is \( x \). The last scenario considered has

\[ f(x) = \phi(x; \mu, \sigma) \]
\[ f(y) = 0.99\phi(y; \mu, \sigma) + 0.01\phi(y; \mu - 10\sigma, \sigma). \]

Again, \( x \) is the better investment. The distributions selected previously are not to be considered real investments. Instead, they were chosen because each case has a clear distinction between the two. For each scenario, we simulate 1000 return streams of sample size 1000 from both distributions. For each return stream, we calculate the risk measures of both investments and count how many times the risk measure correctly determines which investment is most risky. For all risk measures, \( q \) is selected as the mean. For VaR and the accumulate VaR, \( P \) is selected as the lower 5th percentile, unless otherwise noted.

6. Risk measure simulation results

We now present the simulation results and begin with the possibility of a large gain scenario. The percentage of times each measure correctly selected the more risky investment is presented
Table 1. The percentage correct from the possibility of a large gain scenario.

<table>
<thead>
<tr>
<th>Measure</th>
<th>% Correct</th>
</tr>
</thead>
<tbody>
<tr>
<td>Volatility</td>
<td>00</td>
</tr>
<tr>
<td>Coefficient variation</td>
<td>93</td>
</tr>
<tr>
<td>Fishburn $\alpha = 0$</td>
<td>70</td>
</tr>
<tr>
<td>Fishburn $\alpha = 2$</td>
<td>58</td>
</tr>
<tr>
<td>Fishburn $\alpha = \infty$</td>
<td>51</td>
</tr>
<tr>
<td>VaR</td>
<td>54</td>
</tr>
<tr>
<td>Accumulate VaR</td>
<td>51</td>
</tr>
</tbody>
</table>

Note: Notice how the volatility penalizes the gain.

In Table 1, we see as expected that the volatility measure penalizes for dispersion that is created by a gain. Because the coefficient of variation is scaled by the mean, it does a nice job of selecting the correct investment. Also, notice how the other risk measures, which consider only the lower tail, do not distinguish between the two investments. This is interesting because a portfolio manager who sometimes receives a slightly better gain with no additional risk would not be any better than one who receives no additional gain, as decided by these risk measures.

The next scenario is somewhat more realistic in the sense that the returns are skewed towards the losses. This is the moderate chance of a large loss scenario. The results are presented in Table 2. Here, we see the risk measures performing as desired. Because the more risky instrument has a fatter lower tail, the risk measures have an easier time separating the two investments. Here, we see that a portfolio manager who sometimes takes on a large loss will be considered more risky than a manager who does not. This is clearly a good scenario for the risk measures.

The final scenario presents some interesting results. For the small chance of a catastrophic loss, we consider three different values of $P$ for both the VaR and the accumulate VaR. We decided to use the 0.01, 0.05, and the 0.20 lower-tail percentiles. The results are displayed in Table 3. The most striking result is the difference in sensitivity to the choice of $P$ for the VaR and the accumulate VaR. The accumulate VaR selects the correct investment consistently, regardless of the choice of $P$. Notice that VaR does not enjoy this advantage. Because VaR is the standard for many investment banks, it is most unfortunate that it has this property. Also, notice that the Fishburn $\alpha = 0$ measure and the coefficient of variation are not very good at detecting the investment that could have a catastrophic loss.

Table 2. For the moderate chance of a large loss scenario, it is seen that the risk measures will easily distinguish between the two investments.

<table>
<thead>
<tr>
<th>Measure</th>
<th>% Correct</th>
</tr>
</thead>
<tbody>
<tr>
<td>Volatility</td>
<td>100</td>
</tr>
<tr>
<td>Coefficient variation</td>
<td>72</td>
</tr>
<tr>
<td>Fishburn $\alpha = 0$</td>
<td>72</td>
</tr>
<tr>
<td>Fishburn $\alpha = 2$</td>
<td>100</td>
</tr>
<tr>
<td>Fishburn $\alpha = \infty$</td>
<td>100</td>
</tr>
<tr>
<td>VaR</td>
<td>99</td>
</tr>
<tr>
<td>Accumulate VaR</td>
<td>100</td>
</tr>
</tbody>
</table>
7. Conclusions

Altogether, we have demonstrated several points about passive investing and common risk measures through a simple, yet powerful and robust technique. Our first observation is that the market portfolio does not exhibit behavior that would allow it to be truly classified as optimal. This was easily seen by the percentage of random portfolios that outperform the market in the time period from 1970 to 2002. Furthermore, by dividing the study into bull and bear years, we see that active management has an opportunity to provide a valuable service to their client. Specifically, the portfolio manager can greatly enhance a client’s portfolio performance by managing the downside during the bear years, or as we have stated, by offering appropriate risk management. Another point we have made is that volatility and returns are not positively correlated. Our evidence presents that on an yearly basis, returns and volatility can have positive, negative, or zero correlation with returns. The results for the risk measure simulations clearly indicate that no risk measure is appropriate for all situations. We believe that using empirical simulation can help to determine whether a market theory holds, before developing high-level mathematics that describes how the market should behave under such theories. The point here is to look before you leap. Furthermore, we also contend that simulation is an effective means to communicate how complicated mathematical measures behave to people who do not have a quantitative background.

References
