Portfolio Size: Revisited

James Chong, Ph.D.
California State University, Northridge
MacroRisk Analytics

G. Michael Phillips, Ph.D.
California State University, Northridge
MacroRisk Analytics
Portfolio Size: Revisited

Abstract
Using a sophisticated sampling technique, randomly constructed stock portfolios were compared to portfolios using all the underlying population and evaluated using 18 different measures. The randomization included portfolio size and portfolio start date to eliminate timing bias from the analysis. By comparing the 18 statistics across the portfolios, average portfolio sizes to reproduce the population characteristics were computed. The optimal portfolio size depended greatly on the criterion being used to judge the adequacy of diversification.

Keywords:
Portfolios, diversification, Monte Carlo, stocks, financial planning, downside risk, CAPM, simulation, holdings
INTRODUCTION
This paper readdresses an old topic, that of the impact of the number of holdings in a portfolio on the portfolio’s characteristics.

Investors often recognize the trade-off between diversification and the monitoring costs of greater numbers of assets. With the growing number of investible alternatives, investors may be torn between the temptation to buy some of everything, the temptation to buy just what’s new, and the temptation to say “forget it” to the whole thing and buy an index fund.

The conventional wisdom, depending on whose convention, is that somewhere between 10 and 100 stocks, perhaps as many as 300 stocks, are needed to reach portfolio diversification. In this paper, we examine 18 different portfolio attributes using a computationally intensive portfolio resampling technique and find that the number of stocks needed for diversification does vary tremendously depending on which criterion is being measured.

We find that while there are some criteria which would suggest a moderate sized portfolio isn’t well diversified, and consequently that owning individual stocks isn’t optimal with fewer than 70 or more holdings, most of the criteria support the possibility of small to moderate holdings being able to reproduce population performance.

Our paper begins with a review of the portfolio size literature, a discussion of our computational framework, then an introduction to each of the measures that we use to characterize the portfolios. We end with a review of the findings and suggestions for asset managers and investors.

A HISTORY OF THE PROBLEM AND REVIEW OF LITERATURE
While discussing the issue of covariance of securities in a portfolio, Markowitz (1959, pp.111-112) indirectly addressed the topic of portfolio size in relation to portfolio risk. However, most would attribute the first of such studies to Evans and Archer (1968), who demonstrated the marginal benefits derived from increased diversification; that is, the marginal reduction in
portfolio standard deviation by successive increment in stock holdings.\(^1\) They ran a large-scale simulation whereby stocks were randomly selected, and formed into equally-weighted portfolios (also known as naïve diversification). They concluded that there is little diversification benefits beyond holding 10 stocks in a portfolio. On the other hand, Fisher and Lorie (1970) noted that 95\% of achievable reduction in portfolio risk is obtained by holding 32 stocks and 99\%, by holding 128 stocks. Statman (1987), in using an approach reminiscent of the securities market line (SML), concluded that investors hold at least 30 stocks (and at least 40 stocks if no leverage is involved) for a well-diversified stock portfolio. In an update to his earlier study, Statman (2004) found that for the marginal benefit to exceed the marginal cost of diversification, an investor would require at least 300 stocks.

Thus far, the criterion associated with the above studies is portfolio standard deviation. In a departure from this practice, Domain et al. (2007) used a shortfall approach, consistent with the Safety First principle (Roy, 1952), to determine the appropriate number of stocks in a portfolio—it was deemed to be 164 stocks, that would have at most a 1\% chance of underperforming Treasury bonds. In examining the number of mutual funds needed for a diversified mutual fund portfolio, O’Neal (1997) also considered downside risk, and his results, consistent with those of Chong and Phillips (2011a), suggest that holding more than one mutual fund would provide substantial diversification benefits.

**METHODOLOGY**

**Data**

The data we employed for this study are daily closing prices of common stocks listed on the New York Stock Exchange (NYSE) and spans the period January 2, 2003 to November 19, 2010.\(^2\) The period under consideration (as illustrated in Figure 1) covers the Great Recession as well as the preceding and ensuing market run-up and thus, will be a good test for the stability of our results, if any.

---

\(^1\) Analytical solutions between portfolio size and portfolio variance/standard deviation were provided respectively by Elton and Gruber (1977), and Bird and Tipper (1986).

\(^2\) The data was provided by MacroRisk Analytics.
Random portfolio

To create a random portfolio, we perform the following process:

1) For each performance measure (as we’ll describe later), we carried out 1,000 iterations.
2) In each iteration, we
   a) pick a starting date between January 2, 2003 and November 19, 2009;
   b) pull the data from the subsequent year (approximately 250 trading days);
   c) randomly select 100 stocks from the NYSE;
   d) create equally-weighted portfolios comprising 1, 2, 3, …, 100 stocks (totaling 100 portfolios);
   e) compute the statistic of interest.

At the end of the 1,000 iterations (simulations), 100 averages were computed giving the mean value of the performance measure that was being calculated. The 100 averages correspond to the number of portfolio holdings. With this approach, the simulation-based approach avoids possible biasness introduced by the investment process (Kritzman and Price, 2003).

Random starting date. Although our simulation-based approach is similar to that of Evans and Archer (1968), and most studies that followed on this topic, a significant difference is the inclusion of a randomly selected starting date. This, coupled with the extraordinary period covered by our study, would minimize biasedness from the selection of start/end dates and the time horizon considered in this study.

Time diversification. Time diversification “is the idea that the risk of meeting an investment objective can be reduced if risky portfolios with high expected returns are held over long periods of time” (Butler and Domain, 1991). By both increasing the portfolio size and holding the portfolio for long periods of time, equity shortfall risk (Roy, 1952) tends to decrease. However, with regard to time diversification, an exception would arise if idiosyncratic volatility is very high or when correlations are high; then, the holding period may prove ineffective for portfolios

---

3 For discussions on the averaging process employed in simulations, see Newbould and Poon (1993), and Surz and Price (2000).
4 Time diversification is possible because stock returns are not perfectly positively correlated over time.
with a small number of holdings (Milevsky, 2003). As such, by considering portfolio performance over the next year, a fixed length of time, we have ignored the issue of time diversification.

**Maximum number of stocks.** The maximum number of stock holdings used in our simulation is 100. While this may appear wholly inadequate, considering that at least 300 stocks may be needed for diversification (as suggested by Statman, 2004), it is approximately the average number of stocks (102.15) held across 2,294 U.S. domestic equity funds over the period 1980 to 2002 (Jiang et al., 2007).

**Naïve diversification.** Naïve diversification (i.e., the equal investments in all stocks in a portfolio) could reasonably be assumed to be an investor’s preferred strategy (De Wit, 1998; Benartzi and Thaler, 2001) since investors do not have information about future returns (co-)variances (Elton and Gruber, 1977) and are therefore equally likely to choose a particular stock over another (Cohen and Pogue, 1968). Further, in a recent study by DeMiguel et al. (2009), they concluded that, out-of-sample, none of the 14 models under consideration, in terms of the Sharpe ratio, is consistently better than naïve diversification.

**Performance evaluation**

To evaluate how each random portfolio performed against the various performance measures, we use a 100-stock portfolio as benchmark and compare the difference in value of each random portfolio, of incremental stockholding, to the benchmark. For example, taking a particular performance measure, the values of a 50-stock and 51-stock portfolio are respectively $P(50)$ and $P(51)$. The relative performance of $P(51)$ over $P(50)$ is $(P(51) - P(50))/P(100)$, whereby the value for the 100-stock portfolio is $P(100)$. This is consistent with the evaluation methodology employed by Statman (1987), and Newbould and Poon (1993).

---

5 Other studies, e.g., Statman (1987), use 500 stocks as their maximum holding.
6 For an alternative view on naïve diversification in relation to portfolio diversification, see Woerheide and Persson (1993).
DESCRIPTION AND RESULTS OF PERFORMANCE MEASURES

In Table 1, we present the 18 performance measures under consideration, with a brief description of each. We have grouped these measures by category (Panels A to E). While some measures are familiar to readers (or are self-explanatory and would require no explanation), others are less so (or are worth additional discussion), with which we will elaborate.

Previous studies on this research area focus mainly on portfolio risk. In our case, we involved significantly many more performance measures—in addition to those that concerned diversification (categorized as “risk”), we included measures for returns, downside risk, risk-return ratios, and co-movement—to ensure the robustness of our results. Calculating the co-movement measures also provides some insight into the potential number of holdings needed to construct replication portfolios or to match market benchmarks.

Returns

The first two measures (of Table 1, Panel A) are the mean (MeanRet) and median return (MedianRet), which is then followed by the “trimmed mean return.” We trimmed the top and bottom 1% of mean returns (TrimMean), for the reason that the tails contain extreme values which may result in gross measurement errors (Flavin and Yamashita, 2002).

We then consider skewness and kurtosis, which have become increasingly important characteristics in modeling stock returns (Chiang and Li, 2007; Premaratne and Bera, 2005). Positive skewness, in particular, may be desirable to investors. However, by increasing stockholdings and therefore diversification, positive skewness may be diminished. As a result, it has been suggested that some investors should possess a limited number of stockholdings, whereby the exact number is being determined by the investor’s awareness of skewness (Simkowitz and Beedles, 1978; Conine and Tamarkin, 1981).

In Figure 2 below (and subsequent figures), the darker colored line illustrates the particular performance measure for each of the 1,000 random portfolios for each stockholding size, from stockholding of one to 100. The lighter colored line depicts each random portfolio’s performance against the benchmark portfolio of 100 stocks (as described under “Performance evaluation”).

As illustrated in Figure 2, for MeanRet, its relative performance appears to stabilize at about 30 stocks, while those for MedianRet and TrimMean are at about 30 and 40 stocks.
respectively. Skewness registered an optimal stockholding of approximately 30. In the case of Pctile05DailyRet (the 5th percentile of daily returns), its relative measure plateaued at 20 stocks. Kurtosis also required 20 stocks for its relative performance to match that of the 100-stock portfolio.

Risk
Table 1, Panel B covers the various risk measures, the first of which is the portfolio standard deviation (StDev), followed by beta (BetaAvg) of the Capital Asset Pricing Model (CAPM). An empirical specification of the CAPM can be expressed as

$$\left(r_j - r_f\right)_t = \alpha_j + \beta_j (r_m - r_f)_t + \varepsilon_t,$$

where $r_f$ is the observed risk-free rate; $r_j$ is the observed return on asset $j$; $\left(r_j - r_f\right)_t$ is the observed excess return on asset $j$; $\alpha_j$ is the estimated regression intercept, called alpha; $(r_m - r_f)_t$ is the estimated excess return on the market index (here, the S&P 500 index); and $\varepsilon_t$ is the unexplained portion of the model.

The beta’s relative performance to the benchmark portfolio is noisy and stabilized at about 65 stocks (Figure 3). However, the relative performance of portfolio risk, as measured by the standard deviation of returns, suggest that 20 stocks would be sufficient for diversification. While this is twice the number of stocks recommended by Evans and Archer (1968), it is half that suggested by Statman (1987) and substantially fewer than those of other studies.

The “Range” is an alternative measure of returns volatility, which involves the high-low range of portfolio values (Martens and van Dijk, 2007). Following are the standard deviation of the price relative (StDevPRel) and the 5th percentile variance of price relative (VAR05pctPRel).

The price relative is 1 plus the total return from owning a given portfolio. For a given iteration of the simulation, there is a return generated from holding the entire 100-stock sample just as there is a return from each of the other 99 portfolios. The price relative statistics that we compute look at the dispersion of these returns across all 1,000 simulations, holding sample size constant for each calculation. This allows looking at the dispersion of total returns from holding, say, a 30-stock portfolio compared to, say, a 5-stock portfolio or to the whole 100-stock sample.
We compute the range and the standard deviation of the price relative (holding portfolio size constant), and also the variance of the fifth percentile return popularly used for portfolio stress testing (Chong, 2004). These tests help to see the extent to which adding holdings better reproduces the underlying returns distribution. Figure 3 shows that whether one is considering the Range, StDevPRel or VAR05pctPRel, an optimal portfolio size is reached with 30 stocks, a result consistent with central limit theorems (Barbieri et al., 2010).

**Downside risk**

If indeed investors are concerned with capital preservation and loss avoidance, then beta may not be the relevant measure since it considers positive returns in its computation. The correct measure would be one that estimates downside risk (Table 1, Panel C). The first measure of downside risk, in this study, is the semivariance\(^7\) of returns. We approximate semivariance (ASV) with the expression provided by Choobineh and Branting (1986):

\[
ASV_h \approx \left[ P^{1/2}(h - \mu) + (1 - P)^{1/2}\sigma \right]^2.
\]  

(2)

In equation (2), \(h\) is the critical value (e.g., required return on equity) set by the investor, \(\mu\) and \(\sigma\) are respectively the mean and variance of stock returns, and \(P\) is the probability below the critical value.

An alternative to the semivariance is the down-market beta. It resides within the dual-beta model (Chong and Phillips, 2011b), which consists of two regimes, that of an “up-market” and a “down-market”. The former is said to prevail when the market index daily return is non-negative and for the latter, when the market index daily return is negative. The dual-beta equation is similar to (1) except for separate parameters for each regime:

\[
(r_j - r_f)_t = \alpha_j^+ D + \beta_j^+ (r_m^+ - r_f)_t D + \alpha_j^- (1 - D) + \beta_j^- (r_m^- - r_f)_t (1 - D) + \varepsilon_t,
\]

(3)

where \(\alpha_j^+, \beta_j^+, \alpha_j^-\), and \(\beta_j^-\) are the estimated parameters for up-market and down-market days respectively; \(r_m^+ = r_m\) on days the market did not decline and \(r_m^- = r_m\) on days it did; \(D\) is a

\(^7\) See also O’Neal (1997) and Beach (2007).
dummy variable, which takes the value of 1 when the market index daily return is non-negative. Equation (3) collapses to equation (1) when there is no beta asymmetry.

In Figure 4, both ASV and DownBetaAvg reached a state of optimality with about 30 and 40 stocks respectively.

**Risk-return ratios**

We now proceed with risk-return ratios (Table 1, Panel D). The Sharpe ratio (Sharpe, 1966) is one of the most frequently used and relevant performance measures in finance. It measures the excess return per unit standard deviation. In recent times, the Sharpe ratio has been extended to assess the performance of hedge funds. Though it is deemed to be inadequate in evaluating the non-normal return distribution of hedge funds, “research on hedge fund data that compared the Sharpe ratio with other performance measures, however, found virtually identical rank ordering by the various measures” (Eling, 2008). In Figure 5, optimal holding, with the Sharpe ratio, is 25 stocks.

An alternative to the Sharpe ratio is the Sortino ratio (Sortino and Price, 1994), which accounts for downside risk, replacing standard deviation with the square root of the semivariance. In this study, we employed the approximate semivariance (ASV) instead and labeled this measure “QuasiSortino”. As with the Sharpe ratio, QuasiSortino suggests a holding of 25 stocks for reaching an equilibrium state with respect to the Sortino ratio.

**Co-movements**

The last category of measures is related to co-movements (Table 1, Panel E). The Pearson correlation is what investors usually termed “correlation” while the Spearman correlation is its counterpart, using ranked variables. In our study, the Pearson and Spearman correlations are between the portfolio and the “population” returns, where the “population” is the 100-stock portfolio comprising all the assets selected for the iteration. If the smaller portfolios are adequately reproducing the returns of the 100-stock population, then one would expect there to be a high correlation. In Figure 6, the Spearman’s relative measure to the 100-stock portfolio appears to suggest optimal holdings of 45 stocks while, in contrast, Pearson’s is at 35 stocks.

“Coherence”, as we define it, is the percent of times the portfolio changed in the same direction as the S&P 500 index. It reached an equilibrium state fairly quickly with 20 stocks.
APPLICATIONS AND CONCLUSIONS

What becomes clear is that there is generally quick gains to diversification, with (on most measures) a limiting effect taking place at some number of holdings.

The question of whether individual stocks have become more volatile was examined by Campbell et al. (2001) and they concluded the “the number of stocks needed to achieve a given level of diversification has increased.” From our analysis (Table 2), the average number of stocks across the 18 performance measures is 31, far fewer than recommended by recent studies on the topic of portfolio size.

One of the fundamental assumptions underlying optimal portfolio creation and its newer version, portfolio replication, is that one can find portfolios with similar statistical characteristics without needing to buy a mirror of the whole market or the whole population being targeted.

We find that whether one is likely to succeed or not in large part depends on the definition of success.

If the goal is to find a smaller portfolio with betas similar to the 100-stock portfolio, about 65 holdings are required. If one is interested in reproducing the return semivariance or average return, whether mean or median, then 30 stocks are required. A few measures, including coherence with the S&P 500 index and the standard deviation of daily portfolio returns, are able to match the larger portfolio with just 20 stocks. None of our measures converged to the 100-stock portfolio values with fewer than 20 stocks, and only one greater than 50.

A final question is prompted by these results: Why do mutual funds on average have more than 100 holdings, and could funds be effectively replicated with far fewer holdings? While we begin to address this in Chong and Phillips (2011c), a full answer must be left to future research.
REFERENCES


Table 1. List of performance measures.

<table>
<thead>
<tr>
<th>Abbreviations</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Panel A: Returns</strong></td>
<td></td>
</tr>
<tr>
<td>MeanRet</td>
<td>Average (arithmetic mean) return over the holding period.</td>
</tr>
<tr>
<td>MedianRet</td>
<td>Median of the daily returns.</td>
</tr>
<tr>
<td>TrimMean</td>
<td>Average mean return over the holding period, trimming 1% of outliers on each end of the distribution.</td>
</tr>
<tr>
<td>Pctile05DailyRet</td>
<td>5th percentile of daily returns of portfolio.</td>
</tr>
<tr>
<td>Skewness</td>
<td>Skewness of total returns distribution.</td>
</tr>
<tr>
<td>Kurtosis</td>
<td>Kurtosis of total returns distribution.</td>
</tr>
<tr>
<td><strong>Panel B: Risk</strong></td>
<td></td>
</tr>
<tr>
<td>StDev</td>
<td>Standard deviation of the daily returns.</td>
</tr>
<tr>
<td>BetaAvg</td>
<td>Average beta for portfolios of given size.</td>
</tr>
<tr>
<td>Range</td>
<td>The range between the maximum and the minimum values.</td>
</tr>
<tr>
<td>StDevPRel</td>
<td>The standard deviation of the price relative at the end of holding periods.</td>
</tr>
<tr>
<td>VAR05pctPRel</td>
<td>5th percentile variance of price relative.</td>
</tr>
<tr>
<td><strong>Panel C: Downside risk</strong></td>
<td></td>
</tr>
<tr>
<td>ASV</td>
<td>Approximate semivariance of the daily returns.</td>
</tr>
<tr>
<td>DownBetaAvg</td>
<td>Average down market beta for portfolios of given size.</td>
</tr>
<tr>
<td><strong>Panel D: Risk-return ratios</strong></td>
<td></td>
</tr>
<tr>
<td>Sharpe</td>
<td>The Sharpe ratio.</td>
</tr>
<tr>
<td>QuasiSortino</td>
<td>A Sortino ratio, with the square root of the ASV as denominator.</td>
</tr>
<tr>
<td><strong>Panel E: Co-movement</strong></td>
<td></td>
</tr>
<tr>
<td>Pearson</td>
<td>Pearson correlation between portfolio and “population” returns.</td>
</tr>
<tr>
<td>Spearman</td>
<td>Spearman correlation between portfolio and “population” returns.</td>
</tr>
<tr>
<td>Coherence</td>
<td>Percent of times the portfolio changed in the same direction as the S&amp;P 500 index.</td>
</tr>
</tbody>
</table>
Table 2. Optimal number of stocks by performance measure.

<table>
<thead>
<tr>
<th>Abbreviations</th>
<th>Number of stocks</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Panel A: Returns</strong></td>
<td></td>
</tr>
<tr>
<td>MeanRet</td>
<td>30</td>
</tr>
<tr>
<td>MedianRet</td>
<td>30</td>
</tr>
<tr>
<td>TrimMean</td>
<td>40</td>
</tr>
<tr>
<td>Pctile05DailyRet</td>
<td>20</td>
</tr>
<tr>
<td>Skewness</td>
<td>30</td>
</tr>
<tr>
<td>Kurtosis</td>
<td>20</td>
</tr>
<tr>
<td><strong>Panel B: Risk</strong></td>
<td></td>
</tr>
<tr>
<td>StDev</td>
<td>20</td>
</tr>
<tr>
<td>BetaAvg</td>
<td>65</td>
</tr>
<tr>
<td>Range</td>
<td>30</td>
</tr>
<tr>
<td>StDevPRel</td>
<td>30</td>
</tr>
<tr>
<td>VAR05pctPRel</td>
<td>30</td>
</tr>
<tr>
<td><strong>Panel C: Downside risk</strong></td>
<td></td>
</tr>
<tr>
<td>ASV</td>
<td>30</td>
</tr>
<tr>
<td>DownBetaAvg</td>
<td>40</td>
</tr>
<tr>
<td><strong>Panel D: Risk-return ratios</strong></td>
<td></td>
</tr>
<tr>
<td>Sharpe</td>
<td>25</td>
</tr>
<tr>
<td>QuasiSortino</td>
<td>25</td>
</tr>
<tr>
<td><strong>Panel E: Co-movement</strong></td>
<td></td>
</tr>
<tr>
<td>Pearson</td>
<td>35</td>
</tr>
<tr>
<td>Spearman</td>
<td>45</td>
</tr>
<tr>
<td>Coherence</td>
<td>20</td>
</tr>
<tr>
<td>Average stockholdings</td>
<td>31</td>
</tr>
</tbody>
</table>
Figure 1. The S&P 500 index, January 2, 2003 – November 19, 2010.
Figure 2. Performance measure – Returns.
Figure 3. Performance measure – Risk.
Figure 4. Performance measure – Downside risk.
Figure 5. Performance measure – Risk-return ratios.

**Sharpe**

![Sharpe Chart]

**QuasiSortino**

![QuasiSortino Chart]
Figure 6. Performance measure – Co-movement.

Pearson

Spearman

Coherence