

Resampling vs. Shrinkage for Benchmarked Managers

Michael Wolf

Institute for Empirical Research in Economics, University of Zurich, CH-8006
Zurich, Switzerland

Abstract

A well-known pitfall of Markowitz (1952) portfolio optimization is that the sample covariance matrix, which is a critical input, is very erroneous when there are many assets to choose from. If unchecked, this phenomenon skews the optimizer towards extreme weights that tend to perform poorly in the real world. One solution that has been proposed is to shrink the sample covariance matrix by pulling its most extreme elements towards more moderate values. An alternative solution is the resampled efficiency suggested by Michaud (1998). This paper compares shrinkage estimation to

resampled efficiency. In addition, it studies whether the two techniques can be combined to achieve a further improvement. All this is done in the context of an active portfolio manager who aims to outperform a benchmark index and who is evaluated by his realized information ratio.

Keywords

Covariance matrix; Markowitz optimization; Resampling; Shrinkage; Tracking error

1 Introduction

We consider the problem of an active portfolio manager who is measured against the benchmark of an equity market index with fixed (or infrequently rebalanced) weights. In the real world, most managers face a long-only constraint, that is, they are not allowed to short-sell any stock. The common approach to pick stocks in which to invest is mean-variance optimization dating back to Markowitz (1952). It requires two inputs: the expected (excess) return for each stock and the covariance matrix of stock returns. The first input represents the portfolio manager's ability to forecast future price movements. The second input has to be estimated from past stock return data.

The standard statistical method to estimate the covariance matrix of stock returns is to compute the sample covariance matrix. Unfortunately, the sample covariance matrix contains a lot of estimation error. This is especially true when the number of stocks under consideration is large compared to the return history in the sample, which is the typical situation in practice. Feeding the sample covariance matrix to a mean-variance optimizer will result in 'extreme' and under-diversified portfolios. Michaud (1989) calls this phenomenon "error-maximization".

Ledoit and Wolf (2004) propose an improved estimator of the covariance matrix based on the statistical principal of *shrinkage*. The idea is to find an optimal linear combination of the sample covariance matrix and a highly structured estimator, which assumes that the correlation between the returns of any two stocks is always the same. Shrinkage

pulls the most extreme coefficients towards more central values, thereby systematically reducing estimation error where it matters most. An empirical study demonstrates that shrinkage results in a significantly higher realized information ratio of the active manager compared to the sample covariance matrix.

Alternatively, one can apply a variant of the *resampled efficiency* of Michaud (1998) to the problem of the active portfolio manager. This approach is philosophically different from shrinkage estimation, since it is not based upon an improved estimator of the covariance matrix. Instead, one creates artificial return data by *resampling* from the observed data. Then an optimal active portfolio is constructed by computing the sample covariance matrix on the resampled data and feeding it to the mean-variance optimizer. This process is repeated many times and finally the various optimal resampled portfolios are averaged. Typically, the averaged portfolio is more diversified than the one obtained from the sample covariance matrix of the observed data alone. In return, a more diversified portfolio tends to improve out-of-sample performance.

This paper studies the effectiveness of resampling versus shrinkage estimation. In addition, it asks whether by combining resampling with shrinkage one can do better than either technique. We start by giving a formal description of the problem. Next, we briefly recall the shrinkage estimator of Ledoit and Wolf (2004). Thereafter, we describe how to apply resampling to the problem of the active portfolio manager. Finally, we compare the out-of-sample performance of the various methods, using historical stock return data.

2 Formal Description of the Problem

We study the most relevant case for equity portfolios. The benchmark is a weighted index of a large number N of individual stocks, such as a value-weighted index. The universe of stocks from which the portfolio manager selects includes all these stocks.¹ Excess returns are defined relative to the chosen benchmark. Define the following notations:

$$\begin{aligned} w_B &= \text{vector of benchmark weights for the universe of } N \text{ stocks} \\ x &= \text{vector of active weights} \\ w_P &= w_B + x = \text{vector of portfolio weights} \\ y &= \text{vector of stock returns} \\ \mu &= E(y) = \text{vector of expected stock returns} \\ \alpha &= \mu - w_B' \mu = \text{vector of expected stock excess returns} \\ \Sigma &= \text{covariance matrix of stock returns} \end{aligned}$$

We can write expected returns and variances in vector/matrix notation as:

$$\begin{aligned} \mu_B &= w_B' \mu = \text{expected return on benchmark} \\ \sigma_B^2 &= w_B' \Sigma w_B = \text{variance of benchmark return} \\ \mu_P &= w_P' \mu = \text{expected return on portfolio} \\ \sigma_P^2 &= w_P' \Sigma w_P = \text{variance of portfolio return} \\ \mu_E &= x' \mu = \text{expected excess return on portfolio} \\ \sigma_E^2 &= x' \Sigma x = \text{tracking error variance} \end{aligned}$$

The portfolio selection problem is subject to the constraint that the portfolio be fully invested, that is, the portfolio weights w_P have to add up to unity. With $\mathbf{1}$ denoting a conforming vector of ones, this can be written as $w_P' \mathbf{1} = 1$. Because the benchmark weights also add up to unity, the vector of portfolio deviations must sum to zero, or $x' \mathbf{1} = 0$. Therefore, the portfolio of the manager can be viewed as a position in the benchmark plus an *active portfolio*. The active portfolio is a long/short portfolio and expresses the views of the manager. Two immediate implications are:

$$\begin{aligned} \mu_P &= \mu_B + \mu_E \\ \sigma_P^2 &= \sigma_B^2 + 2w_B' \Sigma x + \sigma_E^2 \end{aligned}$$

While positions of the active portfolio are both positive and negative, the manager does not have complete freedom. None of the portfolio weights w_P can be negative, or $w_P \geq 0$, due to the long-only constraint. The resulting constraint $x \geq -w_B$ expresses the limited freedom of the manager. Grinold and Kahn (2000, Chapter 15) illustrate how this limitation can negatively affect the performance of the managed portfolio, especially when the benchmark is a value-weighted index and when N is large. In addition, the manager often is faced with the constraint that the total position in any given stock cannot exceed a certain value, like 10%. If this upper bound is denoted by c , the resulting constraint on the weights of the active portfolio is $x \leq c\mathbf{1} - w_B$.

Having defined the various ingredients, we can now formalize the optimization problem of the manager as follows:²

$$\begin{aligned} \text{Minimize: } & x' \Sigma x \\ \text{such that: } & x' \alpha \geq g \end{aligned}$$

$$\begin{aligned} x' \mathbf{1} &= 0 \\ x &\geq -w_B \\ x &\leq c\mathbf{1} - w_B \end{aligned} \tag{1}$$

Here g is the manager's target gain (i.e., expected excess return) relative to the benchmark. A typical number is 300 basis points (annualized). The manager chooses g and the upper limit c and also knows the current vector of benchmark weights w_B . She is now left to provide estimates for α , the vector of expected stock excess returns, and for Σ , the covariance matrix of stock returns. In a final step, all the inputs are fed into a quadratic optimization software that will compute x , the optimal weights of the active portfolio.

3 Shrinkage Estimator of the Covariance Matrix

Ledoit and Wolf (2004) propose a shrinkage estimator of the covariance matrix, and the reader is referred to their paper for the details. In a nutshell, the estimator is defined as

$$\hat{\Sigma}_{\text{shrink}} = \hat{\delta}^* F + (1 - \hat{\delta}^*) S \tag{2}$$

This equation has three ingredients, S , F , and $\hat{\delta}^*$. S is the sample covariance matrix of the past stock returns. F is a highly structured estimator of the covariance matrix. It assumes that the correlation between the returns of any two stocks is constant. Therefore, F could be coined the constant correlation estimator. $\hat{\delta}^*$ is the shrinkage intensity, as it determines the weight of F in the convex linear combination between S and F . The practical challenge lies in determining $\hat{\delta}^*$ from the data. Ledoit and Wolf (2004) give the formula for that, based on an optimality criterion for the shrinkage estimator.³

4 Resampling

We now describe a variant of the resampled efficiency of Michaud (1998) applied to the problem of the active portfolio manager. In this approach, the manager uses his forecast $\hat{\alpha}$ of future expected excess returns. He employs resampling to avoid solely relying on the sample covariance matrix S computed from past returns. The algorithm is as follows.

Algorithm 4.1 (Resampling the active portfolio)

- Step 1** Resample from the past returns to create a bootstrap sequence of returns.⁴
- Step 2** Compute the sample covariance matrix from the bootstrap data and call it S^* .
- Step 3** Solve the quadratic optimization problem (1) with S^* in place of the unknown true covariance matrix Σ . Call the resulting optimal vector of weights x^* .
- Step 4** Repeat Step 3 K times and average over the K weight vectors x^* to obtain the final vector of active portfolio weights. Call this vector \bar{x}^* .

We want to stress that this algorithm is not identical to the original suggestion of Michaud (1998). Instead of using a 'skilled' forecast,

Michaud also obtains an estimate of the expected excess returns from the bootstrap data.

There are two possibilities for the resampling in Step 1. One can resample from a parametric estimate of the underlying distribution, such as the normal distribution whose mean is the sample mean of the past returns and whose covariance matrix is the sample covariance matrix of the past returns. Alternatively, one can resample from the observed data with replacement. The former corresponds to parametric bootstrap while the latter corresponds to a nonparametric bootstrap; see Efron and Tibshirani (1993). The two approaches yield very similar results in practice.

Is there a general theoretical justification to Algorithm 4.1? To shed some light on this question, it is helpful to realize that the algorithm can be considered a special case of the statistical technique of *bagging*. Bagging is an acronym for “bootstrap aggregating” and was invented by the statistician Leo Breiman (1996). The general situation is as follows. One observes data from which a predictor (or estimator) θ is computed.⁵ But due to the nature of the predictor, small changes in the data set can lead to significant changes in the predictor constructed: the predictor is *unstable*. As a consequence, it is deemed unreliable for practical use. Bagging aims to remedy this situation as follows. One resamples from the observed data via the bootstrap and computes the predictor on the bootstrap data, resulting in θ^* . This process is repeated many times and the resulting values θ^* are ‘aggregated’ by averaging over them. Call the bagged estimator $\bar{\theta}^*$. The hope is that $\bar{\theta}^*$ has a better out-of-sample performance than the original estimator θ . But Breiman (1996) proves there is no universal guarantee. In some applications, $\bar{\theta}^*$ indeed performs better. However, in other applications it is only equally good or even worse than θ . The usefulness of bagging hence must be determined by a case-by-case analysis.

A theoretical investigation of the original Michaud (1998) resampled efficiency is provided by Scherer (2002). In this setting, the portfolio manager aims for an optimal ‘absolute’ portfolio rather than for outperforming an index. In addition, the manager uses the past returns alone to predict future returns as opposed to a ‘skilled’ forecast. Nevertheless, two of the main findings of Scherer (2002) carry over to our modified setting.

First, in the absence of lower and upper bounds on the vector of portfolio weights, resampling is close to the sample covariance matrix. Appendix A details a technical argument. Portfolio selection in the absence of bounds on the portfolio weights is an unstable process. Nevertheless, resampling does not help. So the theoretical analysis reveals an example where the bagged predictor does not outperform the predictor computed from the original data.

Second, in the presence of lower and upper bounds on the vector of portfolio weights, resampling leads to more diversified (active) portfolios compared to the sample covariance matrix. As a result, resampling improves out-of-sample performance. We illustrate the second finding in our context. Assume a specific stock appears very unfavorable according to the sample covariance matrix. The impulse of the mean-variance optimizer is to short-sell the stock. However, due to the overall long-only constraint its weight in the active portfolio cannot be smaller than the negative of its weight in the benchmark index. Suppose then this is the weight finally selected. Now data are resampled from the past returns and the sample covariance matrix is computed and the resulting boot-

strap data. Two things can happen. The first possibility is that, according to the bootstrap matrix, the stock will appear even less favorable than before. But the bootstrap weight cannot be smaller than the original weight due to the overall long-only constraint. So the bootstrap weight will equal the original weight. The second possibility is that, according to the bootstrap matrix, the stock will appear more favorable than before. As a result, the bootstrap weight can be larger than the original weight. These two possibilities imply that the average weight over many bootstrap resamples is likely to be larger (or closer to zero) compared to the original weight. Conversely, a stock that looks very favorable according to the sample covariance matrix is likely to receive a smaller (or closer to zero) weight in the resampled active portfolio. The overall effect is one of a more diversified active portfolio.

To summarize, we agree with Scherer (2002) that resampled efficiency is an heuristic to improve out-of-sample performance in the presence of a long-only constraint but that it is not clear why it should be optimal in any way. On the other hand, shrinkage estimation of the covariance matrix is based on a well-defined optimality criterion.⁶ Moreover, it improves out-of-sample performance whether a long-only constraint is present or not.⁷

5 Resampling Combined With Shrinkage

A natural question to ask is whether resampling can be combined with shrinkage estimation. The answer is yes, as the following algorithm shows.

Algorithm 5.1 (Resampling the active portfolio combined with shrinkage)

- Step 1** Resample from the past returns to create a bootstrap sequence of returns.⁸
- Step 2** Compute the shrinkage estimator (2) from the bootstrap data and call it $\hat{\Sigma}_{Shrink}^*$.
- Step 3** Solve the quadratic optimization problem (1) with $\hat{\Sigma}_{Shrink}^*$ in place of the unknown true covariance matrix Σ . Call the resulting optimal vector of weights x^* .
- Step 4** Repeat Step 3 K times and average over the K weight vectors x^* to obtain the final vector of active portfolio weights. Call this vector \bar{x}^* .

In the presence of a long-only constraint, both resampling and shrinkage estimation improve upon the sample covariance matrix. One might therefore hope that a further improvement can be obtained by combining the two methods. It is not clear how to analyze this question analytically. Instead, we address it in our empirical study.

6 Empirical Study

The set-up is similar to the one of Ledoit and Wolf (2004). We study out-of-sample performance using historical stock market data. DataStream provides monthly U.S. stock data. We use these data to construct several value-weighted indices to serve as our benchmarks. Starting in February 1983, the methodology is as follows. At the beginning of each month, we

select the $N = 30$ largest stocks (with a 10-year history) as measured by their market value. The market values of the stocks define their index weights. At the end of the month, we observe the (real) returns of the individual stocks and, given their weights, compute the return on the index. This prescription is repeated every month until the end of December 2002. Thus, the constituents list and the index weights are constantly updated.

To mimic a skilled active manager, we first construct *raw* forecasts of the expected excess returns by adding random noise to the realized excess returns. In a second step, these *raw* forecasts are transformed into *refined* forecasts $\hat{\alpha}$ which are fed to the quadratic optimizer. This is done in a way such that the unconstrained annualized *ex ante* information ratio (IR) is approximately equal to 1.5, independently of the value of the benchmark size N . The unconstrained IR could be attained by a manager who did not face any lower or upper bound constraints on the weight vector x and who knew the exact nature of the covariance matrix Σ of stock returns. The details of the forecast construction are described in Appendix C of Ledoit and Wolf (2004).

Out-of-sample performance is evaluated in the following way.

Evaluation Algorithm:

- At the beginning of each month feed the following ingredients to the quadratic optimizer: the benchmark weights w_B , the forecasted expected excess returns $\hat{\alpha}$, the estimated covariance matrix $\hat{\Sigma}$, the desired gain g , and an upper bound of $c = 0.1$ on the total weight of any stock.
- For the (annualized) gain g , we use 100, 200, 300, 400, 500, and 600 basis points.
- To compute an estimate $\hat{\Sigma}$, we use the last $T = 60$ monthly returns of the current constituents list of stocks. The quadratic optimizer computes a weight vector x based on $\hat{\Sigma}$.
- If a resampling method is used, apply Algorithm 4.1 or 5.1 based on the last $T = 60$ monthly returns. The resampling method computes a weight vector \bar{x}^* .
- At the end of the month, the realized excess return is given by $e = x'y$ for quadratic optimization and by $e = (\bar{x}^*)'y$ for resampling, where y is the vector of stock returns for the month.
- The out-of-sample period ranges from 02/1983 until 12/2002, so a total of 239 monthly realized excess returns are obtained.
- From the excess returns we compute the (annualized) *ex post* information ratio as $\sqrt{12}\bar{e}/s_e$, where \bar{e} is the sample average of the excess returns and s_e is the sample standard deviation of the excess returns.
- Since the results depend on the monthly forecasts $\hat{\alpha}$, which are random, we repeat this process 50 times and then report mean-summary statistics.

Mean-summary statistics for the realized excess returns are presented in Table 1. The results can be highlighted as follows.

1. In all scenarios, the sample covariance matrix yields the lowest (average) information ratio.
2. Resampling improves upon the sample covariance matrix.
3. Shrinkage improves upon the sample covariance matrix and also upon resampling.

TABLE 1: MEAN-SUMMARY STATISTICS FOR EXCESS RETURNS

	IR	Mean	SD
	gain = 100 b.p.		
Sample	0.89	0.88	0.99
Shrinkage	1.19	1.00	0.84
Resample	0.91	0.88	0.98
Res-Shrink	1.18	0.97	0.84
	gain = 200 b.p.		
Sample	0.95	1.51	1.59
Shrinkage	1.27	1.74	1.38
Resample	0.98	1.53	1.57
Res-Shrink	1.25	1.71	1.37
	gain = 300 b.p.		
Sample	0.97	2.18	2.26
Shrinkage	1.24	2.50	2.03
Resample	1.01	2.23	2.23
Res-Shrink	1.22	2.45	2.02
	gain = 400 b.p.		
Sample	1.00	3.00	3.03
Shrinkage	1.21	3.36	2.80
Resample	1.04	3.10	2.99
Res-Shrink	1.19	3.30	2.80
	gain = 500 b.p.		
Sample	1.04	4.06	3.93
Shrinkage	1.17	4.38	3.76
Resample	1.08	4.16	3.88
Res-Shrink	1.16	4.32	3.75
	gain = 600 b.p.		
Sample	1.09	5.45	5.04
Shrinkage	1.15	5.62	4.92
Resample	1.07	5.23	4.89
Res-Shrink	1.11	5.33	4.80

This table presents *ex post* information ratios, means, and standard deviations of realized excess returns. The out-of-sample period is 02/1983 until 12/2002, yielding 239 monthly excess returns. ‘Sample’ denotes the sample covariance matrix; ‘Shrinkage’ denotes the shrinkage estimator (2); ‘Resample’ denotes Algorithm 4.1; ‘Res-Shrink’ denotes Algorithm 5.1. The results are mean-summary statistics over 50 repetitions.

4. Resampling combined with shrinkage improves upon ‘pure’ resampling but does not improve upon ‘pure’ shrinkage.

Table 2 presents mean-summary statistics on the average monthly turnover. Turnover is defined as the *total turnover* of Grinold and Kahn (2000, Chapter 16) Note that this definition corresponds to updating the entire portfolio, not just the active portfolio. A part of the turnover, therefore, is due to the constituents list of the benchmark and their



TABLE 2: MEAN-SUMMARY STATISTICS FOR AVERAGE MONTHLY TURNOVER

	$g = 100$ b.p.	$g = 200$ b.p.	$g = 300$ b.p.	$g = 400$ b.p.	$g = 500$ b.p.	$g = 600$ b.p.
Sample	0.16	0.28	0.39	0.49	0.58	0.64
Shrinkage	0.12	0.22	0.33	0.44	0.55	0.62
Resample	0.16	0.27	0.38	0.47	0.57	0.62
Res-Shrink	0.12	0.23	0.33	0.44	0.55	0.62

This table presents average monthly turnovers for various strategies. The out-of-sample period is 02/1983 until 12/2002, yielding 239 monthly portfolio updates. ‘Sample’ denotes the sample covariance matrix; ‘Shrinkage’ denotes the shrinkage estimator (2); ‘Resample’ denotes Algorithm 4.1; ‘Res-Shrink’ denotes Algorithm 5.1. The results are mean-summaries over 50 repetitions.

weights, both of which change over time. In general, the turnover is too high to be attractive for an active manager. But no effort was made to limit turnover, and a constraint to this effect could be easily added to the quadratic optimization problem (1). Alternatively, one could address the need to trade by the approach of Michaud (1998) or similar approaches. The important message to take away from Table 2 is that the sample covariance matrix results in the highest turnover, followed by resampling. Shrinkage is comparable to resampling combined with shrinkage.

7 Conclusion

Nobody should be using the sample covariance matrix for the purpose of mean-variance optimization in the context of benchmarked active portfolio management. It places extreme bets on noisy coefficients that contain a lot of estimation error. A superior alternative is the shrinkage estimator of Ledoit and Wolf (2004). It moves noisy coefficients to more common values, resulting in a substantial increase in the realized information ratio of the portfolio manager. A resampling method in the spirit of Michaud (1998) also improves upon the sample covariance matrix. But, according to our empirical study, it is inferior to shrinkage estimation. Resampling can be combined with shrinkage. However, we cannot find evidence that this combination offers any further improvement beyond ‘pure’ shrinkage.

Appendix: Resampling Without Bounds on the Portfolio Weights

Consider the quadratic optimization problem (1) with the upper and lower bounds on the weight vector x removed. The problem then becomes

$$\begin{aligned} \text{Minimize: } & x' \Sigma x \\ \text{such that: } & x' \alpha \geq g \\ & x' \mathbf{1} = 0 \end{aligned} \quad (3)$$

The solution will be a linear combination of the two portfolios $\Sigma^{-1} \mathbf{1}$ and $\Sigma^{-1} \alpha$. More specifically, the optimal portfolio x is given by

$$x = \lambda \Sigma^{-1} \mathbf{1} + \gamma \Sigma^{-1} \alpha$$

where the two constants λ and γ are jointly determined by the two constraints $x' \alpha \geq g$ and $x' \mathbf{1} = 0$. This relation can be proven analogously to equation (4) in Ingersoll (1987, Chapter 4).

In practice, the portfolio manager does not know the vector of expected excess returns α and the covariance matrix of stock returns Σ . Let $\hat{\alpha}$ be her forecast of excess returns and let $\hat{\Sigma}$ be an estimate of the covariance matrix Σ computed from past data. The optimal portfolio x is then

$$x = \lambda \hat{\Sigma}^{-1} \mathbf{1} + \gamma \hat{\Sigma}^{-1} \hat{\alpha} \quad (4)$$

where the two constants λ and γ are jointly determined by the two constraints $x' \hat{\alpha} \geq g$ and $x' \mathbf{1} = 0$.¹⁰ It is easy to see that the optimal solution x does not change if $\hat{\Sigma}$ is replaced by a scalar multiple $c \hat{\Sigma}$ in equation (5), for some constant $c > 0$.

Now apply the resampling Algorithm 4.1 where the quadratic optimization problem (3) replaces the quadratic optimization problem (1). Assume that the resampling is done from a normal distribution with covariance matrix $\hat{\Sigma}$, as suggested by Michaud (1998). If the number of resamples is large, the resampled portfolio \bar{x}^* will be very close to the one determined by the expected value of $(S^*)^{-1}$. Therefore

$$\bar{x}^* \approx \lambda E[(S^*)^{-1}] \mathbf{1} + \gamma E[(S^*)^{-1}] \hat{\alpha} \quad (5)$$

where the two constants λ and γ are jointly determined by the two constraints $x' \hat{\alpha} \geq g$ and $x' \mathbf{1} = 0$. It is well-known that $E[(S^*)^{-1}]$ is a scalar multiple of Σ^{-1} ; for example, see Johnson and Kotz (1972, Section 3). By the above argument, therefore, $\bar{x}^* \approx x$. So the resampled active portfolio will be very close to the Markowitz portfolio.

FOOTNOTES & REFERENCES

1. The problem can be generalized to the setting where the universe contains further stocks not contained in the benchmark. However, to keep transaction costs down, the more general setting is of limited practical interest.
2. Jorion (2003) considers the problem of maximizing $x' \alpha$ subject to an upper bound on the tracking error variance $x' \Sigma x$. Grinold and Kahn (2000) consider the problem of maximizing $x' \alpha - \lambda x' \Sigma x$, where λ is a risk-aversion constant. These are equivalent problem formulations, leading to the same frontier in risk-return space.

3. Computer code in the Matlab programming language implementing this improved estimator is freely downloadable from <http://www.ledoit.net/>.
4. The number of data points in the bootstrap sequence should be the same as in the original sample.
5. In the special case of resampled efficiency, the optimal active portfolio x plays the role of θ .
6. Admittedly, this criterion applies to the estimation of the covariance matrix per se, not the more complex problem of portfolio selection. Nevertheless, an improved estimator of the covariance matrix will result in improved portfolio selection.
7. This constitutes a crucial advantage to a manager who does not face a long-only constraint, such as a hedge fund manager.
8. The number of data points in the bootstrap sequence should be the same as in the original sample.
9. To keep the computational burden feasible, we use $K = 30$ in Algorithms 4.1 and 5.1. We checked for the case $N = 30$ that the results do not change in any meaningful way for larger values of K . This is consistent with Breiman (1996) who finds that most of the practical benefits of bagging are achieved with $K = 25$ already. Both algorithms resample from the observed data with replacement.

- Breiman, L. (1996). Bagging predictors. *Machine Learning*, 24(2):123–140.
- Efron, B. and Tibshirani, R. J. (1993). *An Introduction to the Bootstrap*. Chapman & Hall, New York.

- Grinold, R. C. and Kahn, R. N. (2000). *Active Portfolio Management*. McGraw-Hill, New York, second edition.
- Ingersoll, J. E. (1987). *Theory of Financial Decision Making*. Rowman & Littlefield, Savage, Maryland.
- Johnson, N. L. and Kotz, S. (1972). *Distributions in Statistics: Continuous Multivariate Distributions*. John Wiley, New York.
- Jorion, P. (2003). Portfolio optimization with constraints on tracking error. *Financial Analysts Journal*, 59(5):70–82.
- Ledoit, O. and Wolf, M. (2004). Honey, I shrunk the sample covariance matrix. *Journal of Portfolio Management*, 30(4):110–119.
- Markowitz, H. (1952). Portfolio selection. *Journal of Finance*, 7:77–91.
- Michaud, R. (1989). The Markowitz optimization enigma: Is optimized optimal? *Financial Analysts Journal*, 45:31–42.
- Michaud, R. (1998). *Efficient Asset Management: A Practical Guide to Stock Portfolio Optimization*. Oxford University Press.
- Scherer, B. (2002). Portfolio resampling: Review and critique. *Financial Analysts Journal*, 58(6):98–109.