

## ESTIMATION AND SELECTION BIAS IN MEAN-VARIANCE PORTFOLIO SELECTION

George M. Frankfurter, Christopher G. Lamoureux  
*Louisiana State University*

### Abstract

Much research has focused on the problem of selecting portfolios without the benefit of parametric measures of risk and return. In this paper, a Monte Carlo technique is used to isolate the extent and nature of the problems introduced by this practice. The technique is employed in the context of classical statistical methodology without permitting short sales. It is shown that using estimators of expected return and risk not only obscures parametric values, but also affects portfolio composition in the Markowitz framework. In this study, these two components of bias are isolated and measured.

### I. Introduction

In this paper, the uncertainty in portfolio selection that arises because of imperfect knowledge of the return distributions of risky assets is examined. This risk may be decomposed into two parts: (a) estimation risk; and (b) selection risk. This decomposition makes studying the nature and source of bias in portfolio expected return and risk possible. Further, it is established that some general and ad hoc rules of portfolio diversification can be useful until the normative model of portfolio selection is restructured. Evidence is provided of the inferiority of widely popularized rules of diversification based on normative implications of positive models of asset pricing.

### II. Literature Review

Any application of Normative Portfolio Theory<sup>1</sup> (NPT) involves using estimates of both return and risk, while the original models of Markowitz and Sharpe assume knowledge of the first two moments of the ex-ante return distributions. In practical applications, decision makers do not have access to these parameters. According to statistical theory, estimates of the nondecision (or input) variables of the Markowitz portfolio selection procedure are drawn from a distribution, with the true (but unknown) parameters as their respective means.

Kalymon [13], Barry [2], Brown [4, 5], and Alexander and Resnick [1] employ Bayesian methods to resolve the problem of "estimation risk." Kalymon ([13], p. 560) characterizes estimation risk as a second component of risk: "the decision maker's lack of perfect information about the parameters of his model." In general, the Bayesian School treats the problem by using predictive distributions. Through application

---

<sup>1</sup>NPT refers to the normative model of portfolio selection of Markowitz [16, 17], its simplification by Sharpe [18], and to other works that develop models to deal with normative portfolio decision making. Market equilibrium conditions are not assumed, nor are the normative implications of equilibrium-based models considered.



of these distributions the need for parameter estimates vanishes. Frost and Savarino ([10], p. 303) use a simulation methodology to obtain the result that an " 'all stocks are identical' informative prior selects portfolios whose ex-ante performance is superior to that provided by either the Classical Estimated or the Bayes Diffuse investment rules."

Barry [2], Brown [4], Jobson and Korkie [11], and Klein and Bawa [14] examine the effect of estimation risk on the CAPM. For instance, Jobson and Korkie [11] derive statistical properties of the parameters (e.g., portfolio mean, variance, and weights) in the presence of an orthogonal portfolio, where short selling is unrestricted. Sharpe calls this formulation the "parametric solution," where portfolio weights are dependent on two parameters. In this context, a direct solution of the estimation problem can be found. Unfortunately, no such solution exists under the formulation with non-negativity constraints on portfolio weights, since portfolio weights cannot be derived as a linear combination of a set of parameters.

Thus, a heuristic approach is taken in this paper. In the framework of a Monte Carlo simulation, an understanding is developed of the interactive effects of the selection algorithm's preferences and nominal values of risk-return characteristics.<sup>2</sup> This interaction may induce changes in portfolio composition for small changes in input variables (the vector of means and the variance-covariance matrix in the case of the Markowitz algorithm; alpha, beta, and error variance in the case of the Sharpe [18] simplification of Markowitz). The result of this change in composition may be a major departure from efficiency in parameter space (if the parameters were known) vis-a-vis sample space (when estimates are used).<sup>3</sup>

The contribution of this paper is the separation of estimation bias from selection bias. Through this separation, portfolio managers are better able to gauge portfolio risk-return characteristics, and finance theorists obtain a better understanding of normative portfolio theory.

In the Monte Carlo experiment here, the Sharpe [18] diagonal model is the focus. Frankfurter, Phillips, and Seagle [9] show that the simplified diagonal model tends to perform as well as, or better than, the full variance-covariance model in the presence of limited sample data. Jorion [12] shows that in a full variance-covariance portfolio selection framework, using a Stein-type shrinkage estimator leads to superior portfolio selection. Since the single-index model is used here, the shrinkage estimators and ordinary least squares estimators are isomorphic.

---

<sup>2</sup>Frankfurter, Phillips, and Seagle [8] show that straight-forward application of the Markowitz-Sharpe algorithm will overstate return and understate risk.

<sup>3</sup>The nature of the sensitivity of the portfolio selection algorithm is not addressed directly in this research, although the size of selection bias (as defined herein) is positively related to the sensitivity of the algorithm to perturbations in the inputs. Best and Grauer [3] perform sensitivity analysis in the full variance-covariance context.



### III. Methodology and Data

#### Description

In the Monte Carlo experiment here, an environment is simulated in which asset rates of return are generated by a Gaussian process. Portfolio managers operate under the assumption that the true return-generating process is arithmetic Brownian motion.

To make the simulation relevant, "true" parameters come from estimates obtained with actual data, representing long histories. Monthly NYSE and AMEX CRSP tapes for the period ending December 31, 1985 are used to obtain monthly rate of return data for companies that have two hundred contiguous observations. Accordingly, monthly rates of return are obtained for 680 NYSE stocks and 123 AMEX stocks. Input variables required for the Sharpe model are generated from these data to obtain the parameters of the experiment here.

The CRSP equally weighted NYSE index serves as the common factor in Sharpe's single-index model to obtain OLS estimates. These data are used to estimate:

$$R_{it} = \hat{\alpha}_i + \hat{\beta}_i R_{mt} + \epsilon_{it} \quad (1)$$

Here,  $R_{it}$  and  $R_{mt}$  represent observed (arithmetic) rates of return, in month  $t$  ( $t = 1, 2, \dots, 200$ ) for the stock and index, respectively. OLS estimates of  $\alpha_i$ ,  $\beta_i$ , and  $\sigma^2(\epsilon_i)$ , along with the index mean and variance constitute the parameters of the Monte Carlo experiment. The two key underlying assumptions of the simulation methodology are as follows:

1. Stocks inter-relate with each other solely through their respective "covariances" with the index.
2. The process is stationary.

In this experiment, therefore, neither lack of stationarity, nor incorrect specification of the single-index model can be a source of bias.

Each iteration of the experiment represents a random sample of sixty observations on the index and on each of the 803 stocks. From the index of "parametric" values (mean and variance), a sample of sixty observations on the index is obtained. GGNML (from IMSL) is used to generate the standard unit normal values. Next, from the parametric value of the error variance for each stock, a sample error is generated. These two values ( $\epsilon_{it}$  and  $R_{mt}$ ), together with the 803 sets of  $\alpha_i$  and  $\beta_i$  ("parametric") values, are used to obtain the "sample" observations:

$$R_{it}^j = \alpha_i + \beta_i R_{mt}^j + \epsilon_{it}^j \quad (2)$$

$$j = 1, 2, \dots, 500$$

$$i = 1, 2, \dots, 803$$

$$t = 1, 2, \dots, 60$$



where  $R_{it}^i$  and  $R_{mt}^i$  are rates of return on stock  $i$  and the index, respectively, in simulation  $j$ ;  $\alpha_i$  and  $\beta_i$  are parameters; and  $\epsilon_{it}^j$  is a random error in simulation  $j$  for stock  $i$ . Thus  $500 \times 804 \times 60$  (24,120,000) unique observations are generated.

Next, for each simulation  $j$ , the sixty observations are used to estimate  $\hat{\alpha}_i^j$ ,  $\hat{\beta}_i^j$ ,  $\sigma^2(\epsilon_i^j)$ , and  $E(R_m^j)$  and  $\sigma^2(R_m^j)$  (the mean and variance of the index for iteration  $j$ ) as inputs to the Sharpe simplified model for portfolio selection. Thus, five hundred efficient frontiers (EFs) are obtained. Since each of the five hundred EFs represents a convex set that may map anywhere in mean-standard deviation space, meaningful comparisons require additional theoretical motivation.<sup>4</sup>

### Evaluation

In this research, portfolios are compared according to their shadow prices in parameter space. This approach is used because corner portfolios are unique selections of stocks that totally describe the efficient set, as linear combinations of these corner portfolios provide the entire (continuous) efficient frontier. The shadow price has the following interpretation: the marginal gain, in terms of portfolio expected return, which can be attained by a marginal relaxation of the risk constraint. It also provides a singular measure of preference for a decision maker.

The shadow price of each corner portfolio in parameter space is  $\Theta_\rho^p$ . Since 114 unique corner portfolios are found in parameter space,  $\rho = 1, 2, \dots, 114$ .<sup>5</sup> The higher  $\Theta_\rho^p$  is, the more risk averse the investor is. Confronted with a particular EF, an investor will select a portfolio based on  $\Theta_\rho^p$  (i.e., a reflection of his risk-return preference). Assuming homothetic preferences, an investor will select portfolios, across EFs, with the same  $\Theta$  values. Exploitation of  $\Theta$  in this fashion affords the most general comparisons, in the context of the model. This model is theoretically valid for any expected utility function that depends only on the first two moments of end-of-period wealth. The use of  $\Theta_\rho^p$ , then, for comparison purposes isolates a particular agent's choice— independent of his utility function—when confronted with different choice variables.

### Biases

For the sake of analysis, it is possible to decompose total bias (defined below) into two distinct components. One component is estimation bias. This bias is inescapable, and it results from one's having to make inferences about parameters with limited sample data. The other component is selection bias, a result of portfolio weights being dependent upon sample-estimated values. This bias would not exist with a naive selec-

<sup>4</sup>Five hundred simulations are adequate, as all statistics of interest (e.g., standard deviation of biases) become asymptotic after four hundred simulations.

<sup>5</sup>Only 55 of the 803 stocks (6.85 percent) enter any of the 114 unique corner portfolios, selected from actual market data. There is no constraint on the number of corner portfolios in each simulation. This is also a function of the estimates. Comparing corner portfolios with common shadow prices vitiates any problems that otherwise might arise because of the different nature of the 501 EFs.



tion rule.<sup>6</sup> Thus, the existence of selection bias, vis-a-vis the naive portfolio building rule, may explain the reluctance of portfolio managers and financial theorists to embrace NPT.

In parameter space, corner portfolios represent risk-return combinations that result from multiplying parametric weights by parametric values. These portfolios are unbiased, and thus, truly efficient. In practical application, a decision maker, working with sample values, determines sample weights. Thus, each sample EF is a product of the respective sample weights and estimated values.

Total bias is the difference between what the investor believes the (sample) optimal portfolio parameters are, and the parameters of portfolios that would be selected if all parametric values were known. That is, total bias is the difference:  $(PW \times PV) - (SW \times SV)$ , where  $PW$ ,  $PV$ ,  $SW$ , and  $SV$  stand for parametric weights, parametric values, sample weights, and sample values, respectively. Estimation bias is that part of bias that is the difference between a particular portfolio's (indexed by  $\rho$ ) parametric values and the sample estimators of these parameters:

$$E.B._{\rho} \equiv (SW \times PV) - (SW \times SV) \quad (3)$$

Similarly, selection bias isolates the bias that arises from using sample estimators to select optimal portfolios.

$$S.B._{\rho} \equiv (PW \times PV) - (SW \times PV) \quad (4)$$

It is easy to verify that

$$(PW \times PV) - (SW \times SV) = [(SW \times PV) - (SW \times SV)] + [(PW \times PV) - (SW \times PV)] \quad (5)$$

That is,

$$\text{Total Bias} = \text{Estimation Bias} + \text{Selection Bias}$$

#### IV. Results

Table 1 provides the parametric expected portfolio return,  $E(r_{\rho})^P$ , standard deviation,  $\sigma(r_{\rho})^P$ ,  $\Theta_{\rho}^P$ , and the number of securities in a portfolio,  $k_{\rho}^P$ , for every sixth corner portfolio.<sup>7</sup> The parametric corner portfolios are indexed by  $\rho$  ( $\rho = 1, 2, \dots, 114$ ).

<sup>6</sup>A common norm is to select 16 securities at random. If one accepts the implications (although not all of the assumptions) of the CAPM, then this strategy would be optimal in the sense of no portfolio exposure to diversifiable risk. Such a portfolio will still have estimation bias. If the market does, in fact, conform to the CAPM, then it would be expected that in many simulations virtually all of the 803 stocks would enter at least one efficient portfolio. This is not the case. "Naive" is used to imply a random selection process. In the literature, the most commonly used naive selection method is equal weighting.

<sup>7</sup>In the interest of economy, only every sixth corner portfolio is shown in Table 1. The authors will provide a full list of all corner portfolio values upon request.

TABLE 1. Estimation and Selection Biases (500 Simulations).<sup>a</sup>

$\rho$	$\Theta_\rho^p$	$K_\rho^p$	$E(r_\rho)_\rho^p$	$\sigma(r_\rho)_\rho^p$	$\overline{S.B.}_\rho^m$	$\overline{E.B.}_\rho^m$	$\overline{S.B.}_\rho^\sigma$	$\overline{E.B.}_\rho^\sigma$	$\overline{K}_\rho^s$
6	0.42	6	2.472	8.418	0.954 (0.381) <sup>b</sup>	-4.396 (1.433)	-4.252 (2.697)	0.379 (1.110)	2.842 (1.122)
12	0.77	12	2.387	7.454	0.848 (0.308)	-4.138 (1.366)	-3.328 (2.103)	0.509 (0.930)	4.250 (1.514)
18	1.32	15	2.180	5.925	0.639 (0.254)	-3.805 (1.272)	-3.168 (1.681)	0.627 (0.771)	6.290 (2.020)
24	1.52	18	2.117	5.536	0.579 (0.240)	-3.703 (1.238)	-3.144 (1.551)	0.657 (0.730)	6.960 (2.092)
30	1.83	21	2.030	5.040	0.500 (0.222)	-3.574 (1.197)	-3.149 (1.411)	0.695 (0.686)	7.948 (2.220)
36	2.72	25	1.868	4.253	0.361 (0.184)	-3.237 (1.072)	-2.879 (1.116)	0.785 (0.608)	10.458 (2.652)
42	3.53	28	1.757	3.807	0.273 (0.160)	-3.006 (0.995)	-2.700 (0.975)	0.832 (0.565)	12.426 (2.872)
48	4.12	28	1.696	3.591	0.228 (0.149)	-2.857 (0.956)	-2.558 (0.909)	0.862 (0.543)	13.756 (2.971)
54	5.57	34	1.573	3.210	0.141 (0.129)	-2.561 (0.872)	-2.291 (0.780)	0.906 (0.506)	16.594 (3.279)
60	6.48	36	1.517	3.061	0.104 (0.121)	-2.414 (0.825)	-2.155 (0.715)	0.925 (0.489)	18.044 (3.452)
66	8.00	38	1.445	2.893	0.062 (0.110)	-2.204 (0.774)	-1.939 (0.642)	0.945 (0.464)	20.194 (3.558)
72	8.79	39	1.408	2.814	0.038 (0.107)	-2.110 (0.750)	-1.859 (0.607)	0.953 (0.451)	21.218 (3.647)
78	11.08	38	1.334	2.678	0.000 (0.096)	-1.883 (0.686)	-1.636 (0.517)	0.967 (0.420)	24.096 (4.074)
84	13.37	44	1.288	2.605	-0.017 (0.089)	-1.697 (0.643)	-1.442 (0.460)	0.974 (0.397)	26.308 (4.300)
90	16.69	44	1.236	2.538	-0.036 (0.081)	-1.494 (0.595)	-1.249 (0.391)	0.979 (0.370)	28.780 (4.493)
96	22.69	42	1.183	2.482	-0.046 (0.072)	-1.229 (0.543)	-1.012 (0.315)	0.986 (0.339)	32.294 (4.994)
102	47.75	44	1.110	2.434	-0.040 (0.054)	-0.690 (0.450)	-0.638 (0.179)	0.993 (0.287)	38.870 (6.315)
108	106.55	41	1.076	2.423	-0.026 (0.045)	-0.323 (0.407)	-0.491 (0.117)	0.993 (0.263)	42.018 (7.427)
114	999999	40	1.051	2.421	-0.010 (0.041)	0.024 (0.391)	-0.442 (0.094)	0.992 (0.254)	43.340 (7.658)
MP		(803)	1.127	5.531					

<sup>a</sup>Notes:

$\rho$  = portfolio number (unique corner portfolio along the efficient frontier);

$\Theta_\rho^p$  = shadow price of risk of portfolio  $\rho$  (in parameter space);

$K_\rho^p$  = number of stocks in portfolio  $\rho$  (in parameter space);

$E(r_\rho)_\rho^p$  = expected return on portfolio  $\rho$  (in parameter space);

$\sigma(r_\rho)_\rho^p$  = standard deviation of the returns on portfolio  $\rho$  (in parameter space);

$\overline{S.B.}_\rho^m$  = mean selection bias in the mean of portfolio  $\rho$  (across 500 simulations);

$\overline{E.B.}_\rho^m$  = mean estimation bias in the mean of portfolio  $\rho$  (across 500 simulations);

$\overline{S.B.}_\rho^\sigma$  = mean selection bias in the standard deviation of portfolio  $\rho$  (across 500 simulations);

$\overline{E.B.}_\rho^\sigma$  = mean estimation bias on the standard deviation of portfolio  $\rho$  (across 500 simulations); and

$\overline{K}_\rho^s$  = mean size (in number of securities with nonzero portfolio weights) of portfolio  $\rho$  (across 500 simulations).

<sup>b</sup>Standard deviations (across 500 simulations) are reported in parentheses.



Thus, columns 2, 3, 4, and 5 of Table 1 describe the parametric EF (i.e., the "true" characteristics of the corner portfolios). The mean values of estimation and selection bias, in both the portfolio mean and standard deviation and the mean number of stocks in the portfolio (with their respective standard errors in parentheses below) are presented next. Thus,  $\overline{S.B.}_\rho^m$  represents the mean selection bias (across five hundred simulations) in the portfolio expected return for corner portfolio  $\rho$ .

Figure I is a graphical presentation of Table 1. For additional comparison, the equivalent of a "market portfolio" (MP), as a single point, is also plotted. This is a portfolio that represents the parametric risk and expected return of equal investment in each of the 803 stocks in the universe.

In Figure I, both loci represent the average outcome of five hundred samples. The upper curve represents how the efficient frontier would be constructed, if portfolio managers knew the true parametric values of the universe ( $PW \times PV$ ). Note from Table 1 the consistent, statistically significant overstatement of return (or understatement of risk), at any degree of risk aversion. The bottom curve represents the parametric risk and expected return values of "optimal" portfolios selected from sample data, and portfolio weights that are produced by these data. These weights are then applied to the parametric measures of risk and return. It is clear that internal dominance is present (that is, the curve is not strictly convex). Internal dominance indicates that corner portfolios formed from sample data are dominated by other corner portfolios selected from the same data. This is principally a problem for high-risk portfolios. Thus, portfolio managers would be well advised, as a general rule, to avoid high-risk, high-return portfolios, regardless of the degree of their clients' risk aversion.

It is also clear from the data, and from Figure I, that a naive portfolio (MP) even one with virtually no nonsystematic risk, is inferior to a portfolio that is formed by the Markowitz-Sharpe algorithm. Note that not only does the higher curve provide dominance over MP (a tautology), but so does the lower curve. Thus, the presence of estimation risk should not lead to rejection of the selection model.

The vertical difference between the upper curve in Figure I, and the bottom curve, is the selection bias. This concept helps to explain the complexity of the normative problem. It is interesting to note that this bias is negative, such that its consequence is the reduction of total bias. The absolute selection bias is inversely related to risk. The reduction of this bias is the result of diversification. More diversified portfolios are, *ceteris paribus*, more robust to small perturbations of the sample data. It takes upward of fifteen stocks to adequately exploit the optimization algorithm (since it is being applied to imperfect estimates).

Table 1 reveals that both selection bias and estimation bias decrease with efficient diversification. The behavior of selection bias in the mean is especially interesting. This bias actually changes signs in the middle of the EF. Thus, well-diversified sample corner portfolios may have a parametrically higher expected return than the corresponding parametric corner portfolio. Naturally, though, these corner portfolios have higher risk than the corresponding portfolio (i.e., with the same  $\Theta$ ) in parameter space.

As expected, estimation bias in the portfolio mean tends to shrink in absolute value as diversification is increased along the EF. It is eliminated near the southwest corner of the sample EF. The nature of estimation bias in the standard deviation is

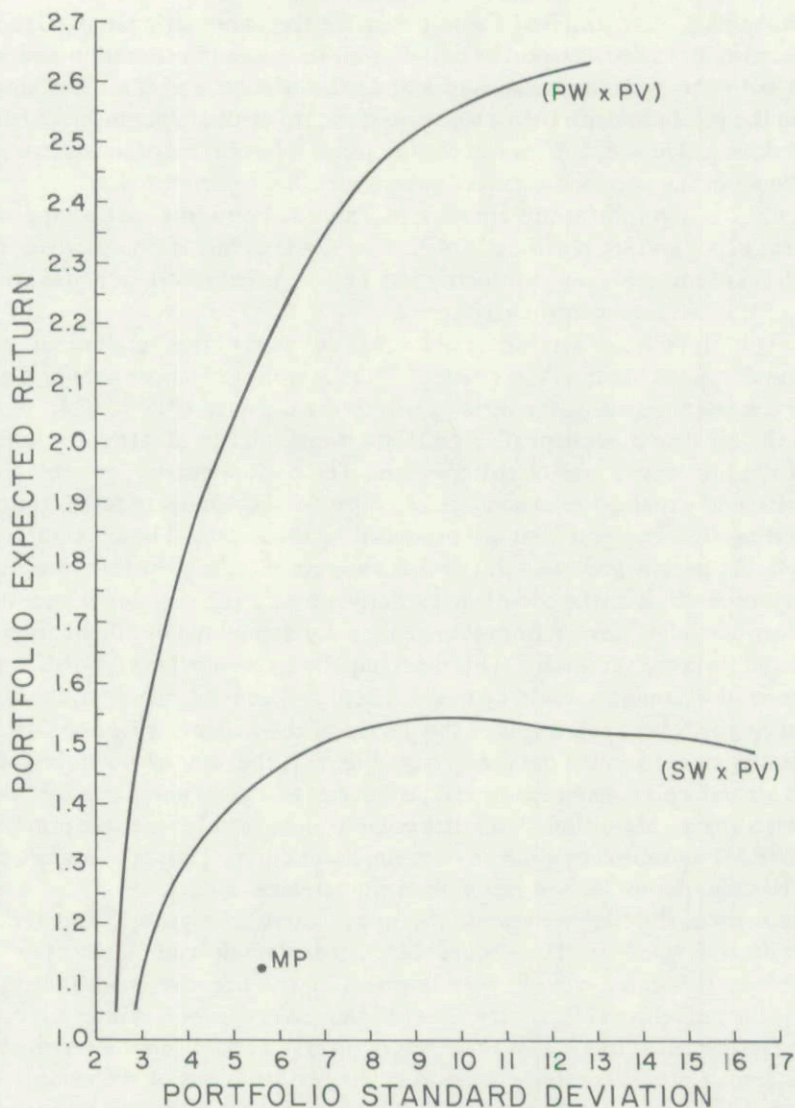


Figure I. The Effect of Estimation Risk on the Mean-Variance Efficient Frontier.

unexpected. The positive signs on this measure in Table 1 indicate that the sample corner portfolio generally lies to the left of the corresponding parametric corner portfolio (based on  $\Theta$ ). This problem is extenuated by increased efficient diversification.

### V. Conclusion

In this paper, portfolio risk in normative applications is decomposed into two distinct parts: risk that results from estimation error, and risk that results from the selection algorithm. Furthermore, the numerical extent of these biases is shown in the



context of a large-scale simulation model; portfolios selected according to the Markowitz-Sharpe maxim are superior to naive selection rules. Interior dominance in the efficient set exists. This dominance is caused by selection bias, and is a result of underdiversification of high-risk, high-expected-return portfolios. As a general rule, optimal portfolios containing at least fifteen stocks are not subject to this interior dominance.

### References

1. Alexander, G., and B. Resnick. "More on Estimation Risk and Simple Rules for Optimal Portfolio Selection." *Journal of Finance* 40 (March 1985), pp. 125-133.
2. Barry, C. "Effects of Uncertain and Nonstationary Parameters Upon Capital Market Equilibrium Conditions." *Journal of Financial and Quantitative Analysis* 13 (September 1978), pp. 419-433.
3. Best, M. J., and R. R. Grauer. "Sensitivity Analysis for Mean Variance Portfolio Problems." Working paper, Simon Fraser University, 1986.
4. Brown, S. "The Effect of Estimation Risk on Capital Market Equilibrium." *Journal of Financial and Quantitative Analysis* 14 (June 1979), pp. 215-220.
5. ———. "Estimation Risk and Optimal Portfolio Choice: The Sharpe Index Model." In *Estimation Risk and Optimal Portfolio Choice*, edited by V. Bawa, C. Brown, and R. Klein, Amsterdam: North-Holland, 1979, pp. 145-171.
6. Elton, E. J., and M. J. Gruber. "Risk Reduction and Portfolio Size: An Analytical Solution." *Journal of Business* (October 1977), pp. 415-437.
7. Evans, J. L., and S. Archer. "Diversification and the Reduction of Dispersion: An Empirical Analysis." *Journal of Finance* 23 (December 1968), pp. 761-769.
8. Frankfurter, G., H. Phillips, and J. Seagle. "Bias in Estimating Portfolio Alpha and Beta Scores." *Review of Economics and Statistics* 56 (August 1974), pp. 412-414.
9. ———. "Performance of the Sharpe Portfolio Selection Model: A Comparison." *Journal of Financial and Quantitative Analysis* 11 (June 1976), pp. 195-204.
10. Frost, P. A., and J. E. Savarino. "An Empirical Bayes Approach to Efficient Portfolio Selection." *Journal of Financial and Quantitative Analysis* 21 (September 1986), pp. 293-305.
11. Jobson, J. D., and B. Korkie. "Estimation for Markowitz Efficient Portfolios." *Journal of the American Statistical Association* 75 (September 1980), pp. 544-554.
12. Jorion, P. "Bayes-Stein Estimation for Portfolio Analysis." *Journal of Financial and Quantitative Analysis* 21 (September 1986), pp. 279-292.
13. Kalymon, B. "Estimation Risk in the Portfolio Selection Model." *Journal of Financial and Quantitative Analysis* 6 (March 1971), pp. 559-582.
14. Klein, R., and V. Bawa. "The Effect of Estimation Risk on Optimal Portfolio Choice." *Journal of Financial Economics* 3 (June 1976), pp. 215-231.
15. ———. "The Effect of Limited Information Risk on Optimal Portfolio Diversification." *Journal of Financial Economics* 5 (August 1977), pp. 89-111.
16. Markowitz, H. "Portfolio Selection." *Journal of Finance* 7 (March 1952), pp. 77-91.
17. ———. *Portfolio Selection: Efficient Diversification of Investments*. New York: John Wiley & Sons, 1959.
18. Sharpe, W. "A Simplified Model for Portfolio Analysis." *Management Science* 9 (January 1963), pp. 277-293.

Copyright of Journal of Financial Research is the property of Blackwell Publishing Limited. The copyright in an individual article may be maintained by the author in certain cases. Content may not be copied or emailed to multiple sites or posted to a listserv without the copyright holder's express written permission. However, users may print, download, or email articles for individual use.