Estimating the Correlation Structure of International Share Prices
Author(s): Cheol S. Eun and Bruce G. Resnick
Source: *The Journal of Finance*, Vol. 39, No. 5 (Dec., 1984), pp. 1311-1324
Published by: Blackwell Publishing for the American Finance Association
Stable URL: http://www.jstor.org/stable/2327729
Accessed: 27/11/2009 07:50

THE JOURNAL OF FINANCE • VOL. XXXIX, NO. 5 • DECEMBER 1984

Estimating the Correlation Structure of International Share Prices

CHEOL S. EUN and BRUCE G. RESNICK*

ABSTRACT

Recently, the case for international portfolio diversification has been convincingly argued in the framework of mean-variance portfolio analysis by a number of researchers. However, virtually no empirical documentation exists concerning the best method for estimating the correlation structure of international share prices. In this paper, 12 models for estimating the international correlation matrix are presented and empirically tested relative to full historical extrapolation. The major evaluation criteria are the mean squared error and stochastic dominance based on the frequency distribution of the squared forecast errors. The results indicate that the National Mean Model strictly dominates all the others in terms of forecasting accuracy.

IN THE LAST FIFTEEN years or so, a number of researchers have convincingly argued the case for international portfolio diversification, as opposed to purely domestic diversification, in the framework of mean-variance portfolio analysis (see, e.g., Grubel [4], Levy and Sarnat [8], Solnik [12], McDonald [10], Lessard [6], and Subrahmanyam [15]).[1] In spite of the emerging consensus in the academic and investment communities that investors can benefit from international diversification in terms of a more efficient risk-return tradeoff, virtually no empirical documentation exists concerning the best method for estimating the correlation structure of international share prices. Since estimating this correlation structure is a necessary precondition for being able to apply modern portfolio theory to international investments, empirical examination is of major importance. Moreover, only through empirical examination of the correlation structure will the effort to devise a simple algorithm that can be used to solve for the optimal international portfolio be facilitated.

In order to trace out the set of ex ante efficient international portfolios using the tenets of modern portfolio theory, the investor must obtain accurate estimates of the expected return and the variance of return for each domestic and foreign security, and he must also obtain accurate estimates of the correlation coefficients of the returns between each pair of securities, stated in terms of his numeraire

* Both authors from Graduate School of Management, University of Minnesota.

[1] With a few exceptions, such as Subrahmanyam [15], most of the studies in support of international diversification have used historical return data to show that internationally diversified portfolios dominate national portfolios in terms of ex post efficiency. As such, they do not offer investigators any direct guidance concerning the best method for constructing ex ante efficient international portfolios.

currency.[2] As was noted by Elton and Gruber [2], for the case of domestic portfolio construction, the task of obtaining accurate estimates of these parameters is particularly difficult with respect to the matrix of correlation coefficients. The reason for this is that the organizational structure, within which the security analysis function is typically performed, does not easily lend itself to extensive inter-firm analysis. Moreover, at the international level, in addition to the usual difficulties of performing domestic security analysis, the analysis of foreign securities is likely to be beset by such difficulties as fluctuating exchange rates, governmental intervention in the capital and foreign exchange markets, varying accounting standards and disclosure requirements across countries, and language barriers. The net effect of this is the scarcity of reliable and comparable information with which to conduct an analysis. Hence, it is highly unlikely that security analysts can generate accurate estimates of correlation coefficients between all pairs of domestic and foreign securities, even if they could do so for domestic securities. Consequently, some type of forecasting model is likely to be the investor's best method for obtaining accurate estimates of the future correlation structure of international share prices.

In this paper, we present a variety of alternative forecasting models for estimating the correlation structure of international share prices. In addition, we present the results from empirically testing and evaluating their respective performances from the viewpoint of an investor who uses the U.S. dollar in measuring asset returns. To insure the robustness of our test results, we used monthly return data for two non-overlapping samples of 80 firms, selected from eight major stock markets and from 12 major industries, from the period of January 1973 through December 1982. This time period roughly corresponds to the period of flexible exchange rates. The first seven years of data (estimation period) were used to generate forecasts of the last three years (forecast period).[3] In evaluating the forecasting models, we are mainly concerned with the question of whether we can confidently discriminate between alternative models on the basis of forecasting accuracy.

The paper is organized as follows. Section I describes the main characteristics of the alternative forecasting models. In Section II, the performance evaluation criteria are discussed. In Section III, the empirical results are presented. A summary and conclusion are offered in Section IV.

I. The Alternative Forecasting Models

In this section, we describe the various forecasting models we employed for estimating the correlation structure of international share prices: a full historical model, three mean models, and eight index models. Since it is difficult, if not

[2] The problems encountered in producing statistical inputs for purely domestic portfolio analysis have been addressed by many authors, e.g., Markowitz [9], Sharpe [11], Cohen and Pogue [1], Elton and Gruber [2], Elton, Gruber, and Urich [3], and Jobson, Korkie, and Ratti [5]. Our paper draws on these previous studies, particularly Elton and Gruber [2] and Elton, Gruber, and Urich [3].

[3] The division of the data into a seven-year estimation period and a three-year forecast period is similar to the proportional split employed by Elton and Gruber [2] and Elton, Gruber, and Urich [3] in their studies of estimating the correlation structure in a domestic setting. Moreover, given the period of flexible exchange rates, it is the most logical split to employ.

Table I

Sample Firms Classified by Country and Industry Membership

Country	No. of Firms	Industry	No. of Firms
1. Australia	10	1. Materials	7
2. France	16	2. Consumer Goods	22
3. Germany	16	3. Automobiles	12
4. Japan	24	4. Banking & Finance	24
5. Netherlands	10	5. Chemicals	16
6. Switzerland	10	6. Electrical & Electronics	18
7. U.K.	24	7. Energy Sources	13
8. U.S.A.	50	8. Health & Personal Care	9
Total	160	9. Merchandising	12
		10. Metals	11
		11. Utilities	6
		12. Machinery & Engineering	10
		Total	160

impossible, to test every conceivable variation of all forecasting models, we choose to test only those models which offer appeal in terms of their logical strength and their underlying economic foundation.

A. Full Historical Model

The most direct and simplest way of estimating the future correlation structure is to calculate each pairwise correlation coefficient over a historical period and assume that the historical values of these coefficients are the best estimates of their future values. In this regard, the future performance of this model would depend on the extent to which history repeats itself. Since each pairwise correlation has to be directly estimated, the Full Historical Model can be regarded as the most disaggregate of all models under consideration. From each of the two non-overlapping samples of 80 international firms, we calculated the 3,160 historical correlation coefficients over the estimation period. Table I shows the distributions of the sample firms by country and industry membership.[4] Throughout the paper, the results for the Full Historical Model serve as the benchmark against which the more aggregate mean models and sophisticated behavioral models are evaluated.

B. Mean Models

An alternative to the Full Historical Model is to hypothesize that the historical correlation matrix contains information about the future mean correlation, but

[4] One hundred sixty sample firms were chosen from among those whose stock prices are quoted for the entire period of January 1973 through December 1982 by *Capital International Perspective.* Reflecting the fact that international investment tends to be concentrated in large firms, in terms of capitalization value, our sample firms tend to be among the largest in their countries. Further, the distribution of sample firms by country tends to reflect the relative size of the national stock markets. Unlike country membership, the industry membership of a firm is not always unambiguous. Therefore, for consistency, we used essentially the classification provided by *Capital International Perspective.* Finally, it is noted that each sample of 80 firms includes the same number of firms from each country and approximately the same number of firms from each industry.

Table II

National Mean Correlation Coefficient Matrix (Estimation Period)

	Australia	France	Germany	Japan	Netherlands	Switzerland	U.K.	U.S.A.
1. Australia	0.5868							
2. France	0.2867	0.5764						
3. Germany	0.1832	0.3121	0.6534					
4. Japan	0.1525	0.2383	0.3006	0.4162				
5. Netherlands	0.2415	0.3447	0.5091	0.2824	0.6245			
6. Switzerland	0.3588	0.3687	0.4751	0.2819	0.5171	0.6640		
7. U.K.	0.3158	0.3782	0.2996	0.2092	0.3934	0.4315	0.6987	
8. U.S.A.	0.3043	0.2259	0.1707	0.1373	0.2714	0.2728	0.2798	0.4392

Data: Various monthly issues of *Capital International Perspective*, Geneva, Switzerland.

not information about pairwise differences from the mean correlation. Mean models assume that the best estimate of the pairwise differences is zero, whereas the Full Historical Model assumes that their historical level is the best estimate. The relative performance of these two types of models, therefore, depends on the degree of intertemporal stability of the pairwise differences.[5]

The most aggregate type of averaging possible is to set every correlation coefficient equal to the average of all (historical) pairwise correlation coefficients worldwide. This model will be referred to as the Global Mean Model. Recognizing, however, that there could be a common mean correlation only within and between subgroups of securities, and also that country and industry membership are the most logical ways of grouping securities, we chose to include two types of disaggregate mean models, namely, the National Mean Model and the Industry Mean Model. For the National Mean Model, every intra-country pairwise correlation coefficient is calculated as being the average of all pairwise correlation coefficients within the country. Further, every inter-country pairwise correlation coefficient between securities from two different countries is calculated as being the average of all inter-country pairwise correlations between securities from the two countries. A corresponding method of averaging is done for the Industry Mean Model.

Table II shows the matrix of the national mean correlations calculated from the entire sample of 160 firms over the estimation period. The intra-country mean correlations range from 0.4162 for Japan to 0.6987 for the U.K., while the inter-country mean correlations range from 0.1373 between the U.S. and Japan to 0.5171 between the Netherlands and Switzerland. As casual observation reveals, the inter-country mean correlations tend to be substantially lower than the intra-country mean correlations. Although not reported here, the inter-industry mean correlations are only slightly lower than the intra-industry mean correlations on the average.

[5] The idea of averaging the historical correlations to estimate their future values was first tested by Elton and Gruber [2], resulting in a surprisingly robust performance.

C. Index Models

Another alternative to direct estimation of the correlation coefficients is to assume that securities move together only because of their common response to a single index or multiple indices. Given a specific behavioral assumption, one can derive the implicit correlation matrix.

C.1 Single-Index Model—The Single-Index Model is described by the following basic equation:

$$R_i = \alpha_i + \beta_i R_W + e_i, \qquad i = 1, \cdots, n \tag{1}$$

with the key assumption that

$$\mathrm{Cov}(e_i, e_j) = 0, \qquad i, j = 1, \cdots, n; \quad i \neq j$$

where R_i and R_W are, respectively, the rates of return on security i and the world market index. This implies that securities are correlated with one another only through their common response to the world market index.[6]

Once β's are estimated via Equation (1), correlation coefficients can be calculated as follows:

$$\rho_{ij} = \beta_i \beta_j \sigma_W^2 / \sigma_i \sigma_j, \qquad i, j = 1, \cdots, n; \quad i \neq j. \tag{2}$$

In our study, we used two different proxies for the world market: a market-value-weighted index and an equally-weighted index of the eight national indices, which are representative of the eight countries from which our two samples were selected. The use of the equally-weighted index was prompted by an earlier finding that it bears a more stable relation to returns on securities and portfolios than a market-value-weighted index.[7]

In addition, we employed two different estimates of β in generating the correlation coefficients: an unadjusted historical β and a Bayesian β. The Bayesian estimation of β involves an adjustment of the sample estimate of β toward the best prior estimate, with the degree of adjustment determined by the precision of both the sample estimate and the prior distribution. As was suggested by Vasicek [17], we used the average value of the sample cross-sectional β's as the best prior estimate.[8] In all, we tested four variants of the Single-Index Model:

i) a historical β with a value-weighted index (SI-HV);
ii) a Bayesian β with a value-weighted index (SI-BV);
iii) a historical β with an equally-weighted index (SI-HE); and
iv) a Bayesian β with an equally-weighted index (SI-BE).

C.2 Multi-Index Model—In contrast to the Single-Index Model, the Multi-Index Model incorporates extra-market indices for the purpose of capturing

[6] Previously, the international version of the Sharpe Single-Index Model was employed by Solnik [13] in testing the International Asset Pricing Model.

[7] See Lessard [7] for a detailed discussion of the advantages of the equally-weighted world index over the value-weighted world index.

[8] See Vasicek [17] for a detailed discussion of this type of adjustment.

additional information. We tested two such models: the Country Index Model and the Industry Index Model. Each model, respectively, purports to capture the country or the industry influence beyond the world market influence.

The Country Index Model tested is of the following specification:

$$R_i = \alpha_i + \beta_i R_W + \gamma_i R_J + \varepsilon_i, \qquad i = 1, \cdots, n \tag{3}$$

with the assumptions that

i) $E(\varepsilon_i) = 0, \qquad i = 1, \cdots, n$
ii) $\mathrm{Cov}(\varepsilon_i, R_W) = 0, \qquad i = 1, \cdots, n$
iii) $\mathrm{Cov}(\varepsilon_i, R_J) = 0, \qquad i = 1, \cdots, n; \quad J = 1, \cdots, N$
iv) $\mathrm{Cov}(R_J, R_W) = 0, \qquad J = 1, \cdots, N$
v) $\mathrm{Cov}(R_J, R_K) = 0, \qquad J, K = 1, \cdots, N; \quad J \neq K$
vi) $\mathrm{Cov}(\varepsilon_i, \varepsilon_j) = 0, \qquad i, j = 1, \cdots, n; \quad i \neq j$

where α_i, β_i, and γ_i are stable parameters specific to security i (of country J) and R_J is the residual return to country J's national market index, constructed to be orthogonal to the world market index. The Country Index Model assumes that a security's returns are affected only by the world market influence and the influence of the particular national market to which the security belongs.[9]

Given the Country Index Model of Equation (3), correlation coefficients can be calculated as follows:

$$\rho_{ij} = \beta_i \beta_j \sigma_W^2 / \sigma_i \sigma_j \tag{4}$$

between firms from different countries, and

$$\rho_{ij} = (\beta_i \beta_j \sigma_W^2 + \gamma_i \gamma_j \sigma_J^2) / \sigma_i \sigma_j \tag{5}$$

between firms from the same country.

The Industry Index Model was also specified in a similar manner, with the only modification being that R_J is now a residual return to industry J's index, constructed to be orthogonal to the world market index. By examination of Equations (2), (4), and (5), it is clear that the Single-Index Model and the Country Index Model will produce identical estimates of inter-country pairwise correlations, but different estimates of intra-country correlations. This, in turn, suggests that any difference in performance between these two types of models should be attributed to their differential estimates of intra-country correlations.

As with the Single-Index Model, we used the two different proxies for the world market index. Thus, altogether, we tested four variants of the Multi-Index Model:

i) a Country Index Model with a value-weighted world market index (MI-CV);
ii) a Country Index Model with an equally-weighted index (MI-CE);
iii) an Industry Index Model with a value-weighted index (MI-IV); and
iv) an Industry Index Model with an equally-weighted index (MI-IE).

[9] Underlying the specification of the Country Index Model in Equation (3) is the assumption that the inclusion of additional foreign market indices is likely to introduce random noise rather than provide additional information. This type of Multi-Index Model has been used by Lessard [7] and Stehle [14].

II. Evaluation Criteria

In evaluating the performance of each of the forecasting models, we employed two different, but complementary criteria: the mean squared forecast error (MSE) and stochastic dominance in terms of the frequency distribution of the forecast errors (SDF). The MSE is perhaps the most popular way of evaluating the quality of a set of forecasts by means of a real number. Given n pairs of corresponding forecasts and actual values (F_i, A_i), the MSE criterion measures the seriousness of forecast errors via the calculation:

$$\text{MSE} = \textstyle\sum_{i=1}^{n} (F_i - A_i)^2/n. \tag{6}$$

In addition, in order to obtain a performance measure in the same dimension as the variable to be forecast, we calculated the square root of the MSE, which is called the root mean squared forecast error (RMS).

To establish the dominance of one model over a second model under the MSE criterion, we first computed the difference, D_i, in the squared forecast errors between each pair of forecasting models for each entry of the correlation matrix:

$$D_i = (F_{i1} - A_i)^2 - (F_{i2} - A_i)^2, \tag{7}$$

where F_{i1} and F_{i2}, respectively, refer to the forecast values of Model 1 and Model 2 for the i^{th} entry of the correlation matrix, and A_i refers to the actual value of the entry. Model 1 is judged to dominate Model 2 if the mean of these differences is negative and significantly different from zero at the 5 percent level. Since each pair of forecasting models produces a "paired" forecast for each entry of the correlation matrix, an ordinary two-tailed T-test can be applied to a "single" mean calculated from the n values of D_i.

To evaluate the forecasting models relative to the benchmark of the Full Historical Model, we computed the Theil Inequality Coefficient (TIC) for each of the forecasting models:

$$\text{TIC} = [\textstyle\sum_{i=1}^{n} (F_i - A_i)^2/(H_i - A_i)^2]^{1/2} \tag{8}$$

where H_i is the value for the i^{th} entry of the Full Historical correlation matrix. It is clear from Equation (8) that TIC > 1 if the forecasts are less accurate than those provided by the historical extrapolation, and TIC = 0 in the event of perfect forecasting.[10]

Under the SDF criterion, we examined the distribution of the differences in the squared forecast errors for the purpose of determining if a model is less likely to make any size error than a second model. Taking a pair of models, one model is said to "dominate" the other if its cumulative frequency function of squared forecast errors, $(F_i - A_i)^2$, is larger than or equal to that of the other at all intervals and strictly larger at one or more intervals. In this study, we examined the frequency distribution at 15 different intervals.[11] In the absence of a strict

[10] For a detailed discussion of the Inequality Coefficient, see Theil [16, p. 28].

[11] Fifteen intervals were taken over the error size range of 0.000 to 0.220, at which the cumulative frequency distribution reaches approximately 96 percent for most of the models. Due to the heavy skewness of the relative frequency distribution toward zero, intervals were taken over increasing distances, i.e., the first four intervals are of 0.005, the second four intervals are of 0.010, the next five intervals are of 0.020, and the last two intervals are of 0.030.

stochastic dominance relationship, one model is said to be "superior" to the other if its cumulative frequency function is larger than that of the other at eight or more intervals out of the 15.

III. Empirical Results

The most significant finding of the empirical tests is that the National Mean Model strictly dominates all the other models under both the MSE and the SDF criteria. As shown in Table III, the National Mean Model has a lower MSE at a statistically significant level than any other model. Further, as shown in Table IV, the probability of producing an error of any size is always smaller with the National Mean Model than with any other model. It is noted that the dominance of the National Mean Model is manifest in the results from both samples. This is in sharp contrast to the rather dismal performances of the two other mean models, the Global Mean and the Industry Mean. Based on this observation, we can deduce the following implications: first, an excessive amount of averaging via the Global Mean Model dilutes the informational content of the historical correlation matrix; second, the dominance of the National Mean Model over the Industry Mean Model confirms the importance of the country factor, which is possibly reinforced by the common currency factor, in determining the dependence structure of international share prices.

Unexpectedly, the second best performing model was the Full Historical Model. It performed significantly better than all the other models in Sample 1 and eight other models in Sample 2 under the MSE criterion. Further, it dominates seven of the other models in Sample 1 and five models in Sample 2 under the SDF criterion, and it is superior to the remaining models. Given the lackluster performance of the Full Historical Model in a domestic setting, it is surprising that it outperformed two mean models (the Industry and the Global), while performing equally well or better than the best of the index models.[12] It is noteworthy that the Full Historical Model performed well in spite of the random noise contained in the historical data. This result probably reflects the model's ability to capture the influence of the country factor.

Despite the fact that both the Full Historical and the Global Mean Models predict the same mean correlation coefficient, the former performed much better. The difference in their performance is attributed to their differential ability to estimate the deviations of the pairwise correlations from the average. The superior performance of the Full Historical Model implies that estimating these deviations at their historical level is better than estimating them as being zero. However, the dominance of the National Mean Model over the Full Historical Model implies that estimating the deviations at their national mean level is better than estimating them at their historical level. To summarize, averaging at the national level is a more efficient way of capturing the informational content of the historical correlations than no averaging at all.

[12] In estimating the dependence structure of U.S. securities, Elton, Gruber, and Urich [3] found that the Full Historical Model was completely dominated by such models as the Overall Mean and the Single-Index.

Table III

Performance and Decomposition of the MSE and Forecast Characteristics of the Alternative Models (Unadjusted)

Forecasting Model	Performance Measures			Decomposition of MSE			Forecast Characteristics		
	MSE[a]	RMS	TIC	U_m	U_v	U_c	$\bar{F}$	σ_F	ρ
Panel A: Results from Sample 1									
1. National Mean	0.0335	0.1830	0.9337	0.1379	0.1704	0.6917	0.3207	0.1265	0.5458
2. Full Historical	0.0384	0.1960	1.0000	0.1203	0.0356	0.8441	0.3207	0.1649	0.5129
3. MI-CE	0.0401	0.2003	1.0219	0.2152	0.1145	0.6703	0.3457	0.1342	0.5032
4. SI-BE	0.0416	0.2039	1.0403	0.0721	0.2882	0.6397	0.3075	0.0927	0.2877
5. MI-CV	0.0424	0.2060	1.0510	0.1440	0.1403	0.7157	0.3309	0.1249	0.3969
6. MI-IE	0.0432	0.2078	1.0602	0.0777	0.1899	0.7324	0.3106	0.1114	0.2966
7. SI-HE	0.0437	0.2090	1.0663	0.0682	0.1926	0.7392	0.3073	0.1105	0.2744
8. Industry Mean	0.0447	0.2114	1.0786	0.1032	0.4834	0.4134	0.3207	0.0548	0.1638
9. SI-BV	0.0448	0.2117	1.0801	0.0488	0.3708	0.5804	0.2995	0.0728	0.1167
10. Global Mean	0.0454	0.2131	1.0872	0.1018	0.8982	0.0000	0.3207	0.0000	0.0000
11. MI-IV	0.0458	0.2139	1.0913	0.0608	0.2795	0.6597	0.3055	0.0889	0.1578
12. SI-HV	0.0464	0.2155	1.0995	0.0508	0.2870	0.6622	0.3013	0.0866	0.1191
Panel B: Results from Sample 2									
1. National Mean	0.0357	0.1889	0.9069	0.1122	0.1692	0.7186	0.2895	0.1244	0.4898
2. Full Historical	0.0434	0.2083	1.0000	0.0923	0.0295	0.8782	0.2895	0.1663	0.4332
3. MI-CE	0.0441	0.2101	1.0086	0.1995	0.1303	0.6702	0.3200	0.1262	0.4202
4. SI-BE	0.0431	0.2077	0.9971	0.0694	0.3159	0.6147	0.2809	0.0853	0.2307
5. MI-CV	0.0473	0.2176	1.0446	0.1196	0.1451	0.7353	0.3014	0.1192	0.2776
6. MI-IE	0.0445	0.2109	1.0125	0.0722	0.2261	0.7017	0.2829	0.1018	0.2409
7. SI-HE	0.0451	0.2124	1.0197	0.0655	0.2221	0.7124	0.2806	0.1019	0.2197
8. Industry Mean	0.0447	0.2115	1.0154	0.0894	0.4698	0.4408	0.2895	0.0571	0.1449
9. SI-BV	0.0472	0.2174	1.0437	0.0421	0.3655	0.5924	0.2708	0.0706	0.0192
10. Global Mean	0.0448	0.2117	1.0163	0.0893	0.9107	0.0000	0.2895	0.0000	0.0000
11. MI-IV	0.0485	0.2203	1.0576	0.0486	0.2837	0.6677	0.2747	0.0847	0.0541
12. SI-HV	0.0495	0.2225	1.0682	0.0414	0.2877	0.6709	0.2715	0.0827	0.0064

[a] All differences in the MSE are statistically significant at the 5 percent level, except among the following groups of models: (4, 5), (5, 6, 7), (7, 8, 9, 10, 11), and (8, 9, 10, 11, 12) in Panel A, and (2, 3, 4), (3, 6, 7, 8, 10), (5, 9), and (5, 11) in Panel B.

Table IV

Dominance by Cumulative Frequency Distribution of the Squared Forecast Errors[a] (Unadjusted)

	National Mean	Full Historical	MI-CE	SI-BE	MI-CV	MI-IE	SI-HE	SI-BV	MI-IV	Industry Mean	SI-HV	Global Mean	Total
Panel A: Results from Sample 1													
1. National Mean		15	15	15	15	15	15	15	15	15	15	15	165
2. Full Historical	0		13	14	15	14	15	15	15	15	15	15	146
3. MI-CE	0	2		13½	14	15	15	15	15	15	15	15	134½
4. SI-BE	0	1	1½		10	13	13½	15	14	15	14½	15	$111(½)^3$
5. MI-CV	0	0	1	5		12	13½	15	15	15	15	14	105½
6. MI-IE	0	1	0	2	3		11	13	15	15	14	15	89
7. SI-HE	0	0	0	1½	1½	4		13	14	15	14	13	$75(½)^2$
8. SI-BV	0	0	0	0	0	2	2		12	10	12	15	53
9. MI-IV	0	0	0	1	0	0	1	3		6	14	11	36
10. Industry Mean	0	0	0	0	0	0	0	5	9		9	8	31
11. SI-HV	0	0	0	½	0	1	1	3	1	6		11	23½
12. Global Mean	0	0	0	0	1	0	2	0	4	7	4		18
Panel B: Results from Sample 2													
1. National Mean		15	15	15	15	15	15	15	15	15	15	15	165
2. Full Historical	0		15	12	15	14	14	15	15	12	15	12	139
3. MI-CE	0	0		9	15	10½	11½	15	15	10	15	10	$110(½)^2$
4. SI-BE	0	3	6		14	12½	15	15	15	13	15	15	123½
5. MI-CV	0	0	0	1		1	3	10	15	8	15	8	61
6. MI-IE	0	1	4½	2½	14		12½	15	15	10	15	11	$99(½)^3$
7. SI-HE	0	1	3½	0	12	2½		15	14	9½	15	10	$81(½)^3$
8. SI-BV	0	0	0	0	5	0	0		9	1	11½	3	29½
9. MI-IV	0	0	0	0	0	0	1	6		2	15	3	27
10. Industry Mean	0	3	5	2	7	5	5½	14	13		14	11	79½
11. SI-HV	0	0	0	0	0	0	0	3½	0	1		3	7½
12. Global Mean	0	3	5	0	7	4	5	12	12	4	12		64

[a] A number in the table represents the number of times, out of 15 intervals examined, that the left-hand side model has a larger cumulative frequency function, i.e., smaller errors, than the model at the top. In cases where two models have the same cumulative frequency function once or n times, out of 15 intervals, that is indicated by ½ or $(½)^n$, respectively. To facilitate multi-lateral comparison, the sum of these numbers is provided for each model in the last column.

The best among the index models were the Country Index Model with an equally-weighted world market index (MI-CE) and the Single-Index Model using a Bayesian beta estimated from an equally-weighted world market index (SI-BE). As shown in Table III, MI-CE has the lowest MSE in Sample 1, while SI-BE exhibits the lowest MSE among the index models in Sample 2. The strong performance of these two models is also corroborated by Table IV. From examining the performance of the four variants of the Multi-Index Model, it is confirmed that the country factor exerts a much stronger influence on the return generating process than the industry factor, and also that the equally-weighted world market index is a better proxy for the world factor than the value-weighted index. The stronger performance of the Full Historical Model over the index models, however, suggests that estimating the correlations between shares at their historical levels, net of their common correlation with the index or indices, is better than estimating them at zero.

As was shown by Theil, the MSE can be decomposed into three terms, each of which represents a particular kind of forecast error:[13]

$$\text{MSE} = \textstyle\sum_{i=1}^{n} (F_i - A_i)^2/n = (\bar{F} - \bar{A})^2 + (\sigma_F - \sigma_A)^2 + 2(1 - \rho)\sigma_F\sigma_A, \tag{9}$$

where $(\bar{F}, \bar{A})$ and (σ_F, σ_A) are, respectively, the means and the standard deviations of the forecast and the actual correlations, and ρ is the correlation coefficient between the forecast and the actual correlations. The first term of Equation (9) measures the error due to a biased forecast; the second term measures the error due to unequal variation; and the third term measures the error due to an imperfect correlation. Dividing through the right-hand side of Equation (9) by the MSE results in three "inequality proportions," which measure the relative importance of the components. These three terms are, respectively, labeled:

$$\text{Bias proportion} \quad U_m = (\bar{F} - \bar{A})^2/\text{MSE}, \tag{10a}$$

$$\text{Variance proportion} \quad U_v = (\sigma_F - \sigma_A)^2/\text{MSE}, \tag{10b}$$

$$\text{Covariance proportion} \quad U_c = 2(1 - \rho)\sigma_F\sigma_A/\text{MSE}. \tag{10c}$$

Obviously, the three proportions add to unity. An immediate implication of the preceding analysis is that once the MSE criterion is accepted, the character of a forecasting model is fully specified by three parameters: $\bar{F}$, σ_F, and ρ. The values of these parameters were computed for each of the forecasting models, and they are presented in the third section of Table III.[14]

The second section of Table III provides the inequality proportions for each of the forecasting models. With the exceptions of the Global Mean Model and the Industry Mean Model, which are the two most aggregate forecasting models, imperfect covariance, as measured by U_c, is by far the largest source of error. For the ten other models, U_c ranges from 0.5804 (0.5924) for SI-BV to 0.8441 (0.8782)

[13] For a detailed discussion of the decomposition of the MSE, see Theil [16, p. 29].

[14] The standard deviation of forecast correlations, σ_F, in fact, reflects each model's ability to differentiate individual pairwise correlations from the mean, $\bar{F}$. The higher σ_F is, the more disaggregate is the forecast that the model provides.

for the Full Historical Model in Sample 1 (Sample 2).[15] The average U_c is 0.6935 for Sample 1 and 0.6962 for Sample 2. The average values of U_m and U_v are 0.0996 and 0.2069, respectively, in Sample 1, and 0.0863 and 0.2175 in Sample 2. Noting the calculation of U_c in Equation (10c), it is not surprising that three of the best performing models (National Mean, Full Historical, and MI-CE) are those with the highest ρ values.

As an additional method of testing, we examined the performance when each model was adjusted to have the same mean forecast. Elton, Gruber, and Urich [3, p. 1378] noted that this method of performance evaluation is useful in determining whether the forecasting accuracy of a model is dependent upon its forecast of the mean correlation. As can be seen from Table III, the mean forecast correlations range from 0.2995 (0.2708) for SI-BV to 0.3457 (0.3200) for MI-CE in Sample 1 (Sample 2). Considering that the mean of the actual correlations is 0.2527 for Sample 1 and 0.2262 for Sample 2, all of the forecasting models overestimated the mean correlation. But some models did so more than the others. In general, the Multi-Index Models tend to produce higher estimates of the mean correlation than the Single-Index Models, while the Full Historical and the mean models forecast the same mean correlation as the average of the historical correlations.

Table V presents the performance results of the forecasting models when the forecasts of each model are adjusted to have the same mean as the average of the historical correlations. This table is the direct counterpart to the unadjusted results presented in Table III. Table V indicates that the National Mean Model continues to outperform all the other models. But the second best performing model turned out to be MI-CE, rather than the Full Historical Model. This was true under both the MSE and the SDF criteria.[16] The stronger performance of MI-CE is due to the fact that the bias in forecasting was significantly reduced as a result of the mean adjustment. This is evidenced by the bias proportion, U_m, decreasing from 0.2152 (0.1995) to 0.1279 (0.1018) in Sample 1 (Sample 2). In general, the Country Index Models improved their ranking in terms of the MSE, while the Industry Index Models and the Single-Index Models experienced deterioration in their ranking. It is pointed out, however, that no model experienced a change in ranking, upward or downward, by more than two places. This is due to the fact that bias does not account for very much of the overall error as compared with the other sources of forecast error.[17]

[15] The U_c is zero for the Global Mean Model because it offers a uniform forecast for each and every pairwise correlation; consequently, it will have an extremely high value for U_v. The opposite is true of the most disaggregate model, the Full Historical; it has a very high U_c value, whereas U_v is near zero.

[16] For economy of space, the detailed performance results under the SDF criterion are not reported here. We obtained similar results to those reported in Table IV, except that with the mean adjustment, MI-CE strictly dominates the Full Historical Model.

[17] In order to determine if the performance results presented in this section are dependent upon the MSE criterion, we also computed the mean absolute forecast error (MAE) for each of the forecasting models. The salient feature of the results is that the National Mean Model continues to dominate all of the other models, with or without the mean adjustment.

Table V

Performance and Decomposition of the MSE of the Alternative Forecasting Models (Adjusted)

	Performance Measures			Decomposition of MSE		
Forecasting Model	MSE[a]	RMS	TIC	U_m	U_v	U_c
Panel A: Results from Sample 1						
1. National Mean	0.0335	0.1830	0.9337	0.1379	0.1704	0.6917
2. MI-CE	0.0361	0.1900	0.9694	0.1279	0.1272	0.7449
3. Full Historical	0.0384	0.1960	1.0000	0.1203	0.0356	0.8441
4. MI-CV	0.0410	0.2025	1.0332	0.1129	0.1454	0.7417
5. SI-BE	0.0432	0.2078	1.0602	0.1070	0.2773	0.6157
6. MI-IE	0.0445	0.2110	1.0765	0.1040	0.1845	0.7115
7. Industry Mean	0.0448	0.2117	1.0801	0.1032	0.4834	0.4134
8. SI-HE	0.0453	0.2128	1.0857	0.1020	0.1856	0.7124
9. Global Mean	0.0454	0.2131	1.0872	0.1018	0.8982	0.0000
10. SI-BV	0.0473	0.2175	1.1097	0.0978	0.3517	0.5505
11. MI-IV	0.0476	0.2182	1.1133	0.0971	0.2687	0.6342
12. SI-HV	0.0487	0.2207	1.1260	0.0949	0.2737	0.6314
Panel B: Results from Sample 2						
1. National Mean	0.0357	0.1889	0.9069	0.1122	0.1692	0.7186
2. MI-CE	0.0393	0.1983	0.9520	0.1018	0.1462	0.7520
3. Full Historical	0.0434	0.2083	1.0000	0.0923	0.0295	0.8783
4. MI-CV	0.0457	0.2137	1.0259	0.0876	0.1503	0.7620
5. SI-BE	0.0442	0.2102	1.0091	0.0906	0.3086	0.6007
6. MI-IE	0.0453	0.2128	1.0216	0.0884	0.2221	0.6895
7. Industry Mean	0.0447	0.2115	1.0154	0.0894	0.4698	0.4408
8. SI-HE	0.0462	0.2149	1.0317	0.0867	0.2171	0.6962
9. Global Mean	0.0448	0.2117	1.0163	0.0893	0.9107	0.0000
10. SI-BV	0.0493	0.2220	1.0658	0.0812	0.3506	0.5682
11. MI-IV	0.0502	0.2240	1.0754	0.0798	0.2744	0.6458
12. SI-HV	0.0515	0.2269	1.0893	0.0778	0.2768	0.6454

[a] All differences in the MSE are statistically significant at the 5 percent level, except among the following groups of models: (6, 7, 8, 9) and (10, 11) in Panel A, and (3, 5, 7, 9), (4, 6, 9), and (4, 8, 9) in Panel B.

IV. Concluding Remarks

In view of the recent trend toward a greater integration of capital markets, as well as a greater awareness on the part of individual and institutional investors of the potential gains from international diversification, it is a necessity to make modern portfolio theory amenable to implementation in an international setting. The purpose of this paper was to make a contribution to that end. We have evaluated the performance of 12 alternative forecasting models that can be used to estimate the correlation structure of international share prices. Because full historical extrapolation is the most disaggregate method of estimation, the results from the Full Historical Model were used as the benchmark against which the other 11 models were evaluated.

The most important result emerging from our empirical tests is that the National Mean Model strictly dominates all the other models in terms of forecasting accuracy. This result corresponds with earlier findings that there is a strong country factor influencing the return generating process. This finding should help other researchers to devise a simple algorithm that can be used to solve for the optimal international portfolio. It should be pointed out, however, that because we have not tested every possible variant of all forecasting models, the National Mean Model may not necessarily be the best model. In view of the finding that "imperfect correlation" is by far the largest source of the overall forecast error, the challenge to future researchers lies in developing a model which is characterized by a higher correlation between the forecast and the actual correlations.

REFERENCES

1. Kalman Cohen and Gerald Pogue. "An Empirical Evaluation of Alternative Portfolio Selection Models." *Journal of Business* 46 (April 1967), 166–93.
2. Edwin Elton and Martin Gruber. "Estimating the Dependence Structure of Share Prices—Implications for Portfolio Selection." *Journal of Finance* 28 (December 1973), 1203–32.
3. Edwin Elton, Martin Gruber, and Thomas Urich. "Are Betas Best?" *Journal of Finance* 33 (December 1978), 1375–84.
4. Herbert Grubel. "Internationally Diversified Portfolios: Welfare Gains and Capital Flows." *American Economic Review* 58 (December 1968), 1299–1314.
5. J. Jobson, B. Korkie, and V. Ratti. "Improved Estimation for Markowitz Portfolios Using James-Stein Type Estimators." *Proceedings of the American Statistical Association, Business and Economic Statistics Section* (1979), 279–84.
6. Donald Lessard. "International Portfolio Diversification: A Multivariate Analysis for a Group of Latin American Countries." *Journal of Finance* 28 (June 1973), 619–33.
7. ———. "World, Country, and Industry Relationships in Equity Returns: Implications for Risk Reduction through International Diversification." *Financial Analyst Journal* 32 (January/February 1976), 32–38.
8. Haim Levy and Marshall Sarnat. "International Diversification of Investment Portfolios." *American Economic Review* 60 (September 1970), 668–75.
9. Harry Markowitz. *Portfolio Selection.* New York: John Wiley & Sons, 1959.
10. John McDonald. "French Mutual Fund Performance: Evaluation of Internationally Diversified Portfolios." *Journal of Finance* 28 (December 1973), 1161–80.
11. William Sharpe. "A Simplified Model for Portfolio Analysis." *Management Science* 9 (January 1963), 277–93.
12. Bruno Solnik. "Why Not Diversify Internationally?" *Financial Analyst Journal* 30 (July/August 1974), 48–54.
13. ———. "An International Market Model of Security Price Behavior." *Journal of Financial and Quantitative Analysis* 9 (September 1974), 537–54.
14. Richard Stehle. "An Empirical Test of the Alternative Hypotheses of National and International Pricing of Risky Assets." *Journal of Finance* 32 (May 1977), 493–502.
15. Marti Subrahmanyam. "On the Optimality of International Capital Market Integration." *Journal of Financial Economics* 2 (March 1975), 3–28.
16. Henri Theil. *Applied Economic Forecasting.* Amsterdam: North-Holland, 1971.
17. Oldrich Vasicek. "A Note on Using Cross-Sectional Information in Bayesian Estimation of Security Betas." *Journal of Finance* 28 (December 1973), 1233–39.