# ON THE ASSESSMENT OF RETURN GENERATING MODELS

by

Marshall E. Blume
Mustafa N. Gultekin
N. Bulent Gultekin

# 13-86

RODNEY L. WHITE CENTER FOR FINANCIAL RESEARCH
The Wharton School
University of Pennsylvania
Philadelphia, PA 19104

# ON THE ASSESSMENT OF RETURN GENERATING MODELS

by

Marshall E. Blume, Mustafa N. Gultekin, and N. Bulent Gultekin*

*The Wharton School, University of Pennsylvania; University of North Carolina at Chapel Hill; and both the Central Bank of the Republic of Turkey and the Wharton School, University of Pennsylvania, respectively. The views expressed here are solely those of the authors and do not necessarily represent the views of the Central Bank. The authors would like to thank Craig MacKinlay for his careful and thoughtful comments.

# ON THE ASSESSMENT OF RETURN GENERATING MODELS

Multifactor asset pricing models and the return generating models that characterize them have become an important research area in financial economics. The Arbitrage Pricing Theory of Ross [22] and the empirical tests of this model have motivated much of the recent interest in this area.

Factor analysis and similar factor analytic techniques have on occasion played an important role in the analysis of returns on common stocks and other types of financial assets. Farrar [14] may have been the first to use factor analysis in conjunction with principal component analysis to assign securities into homogeneous correlation groups. King [18] used factor analysis to evaluate the role of market and industry factors in explaining stock returns. These two studies sparked an interest in multi-index models, and a rich body of empirical work soon emerged. Examples include Elton and Gruber [11, 12], Meyers [20], Farrell [15] and Livingston [19], among others. The major goal in these earlier studies was to establish the smallest number of "indexes" needed to construct efficient sets.

Recent research, however, uses factor analytic techniques with an entirely different orientation. Its aim is to investigate the cross-sectional risk-return relation implied by the Arbitrage Pricing Theory or its variants, commonly called arbitrage pricing models (see for example Conner [8]). Empirical tests of the APT model can be found in Roll and Ross [21], Chen [5], Brown and Weinstein [3], Dhrymes, Friend and Gultekin [9, 10], Gultekin and Rogalski [17], and Cho, Elton, and Gruber [7], to cite a few from the rapidly growing literature.

Most multifactor models, such as the APT, specify neither the minimum number of factors required to generate security returns nor any economic hypothesis as to what the factors should represent.[1] As a consequence, cross-sectional pricing tests of the arbitrage pricing models based upon factor analysis must make an arbitrary decision as to the number of factors to extract or allow the data to determine the number. While there is a growing body of new and interesting work in multifactor asset pricing models, there has been no direct analysis[2] of the properties of the return generating factors based upon factor analysis. The purpose of this paper is to address this gap in the literature by examining the robustness of multifactor return generating models estimated by factor analysis. Specifically, the paper investigates the predictive power of such models and then compares these models to simpler return generating models that are based on market and industry indexes.[3] Our approach differs sharply from the other studies in this area. In this study, we concentrate on the accuracy of the return predictions rather than the accuracy of prediction of correlation matrices of the earlier studies or the more recent significance tests of the cross-sectional pricing relation implied by the arbitrage pricing models.

We demonstrate how the factors in a factor analytic model can be transcribed into standard types of indexes that can be interpreted as returns on specific portfolios. In the process, we show that the factor analysis, as it is often used, tends to point to more factors than are actually present. Moreover, due to the orthogonality requirement of factor analysis, the security weights in the portfolios that comprise the factor scores are such that it is highly unlikely that these portfolios and, thus, the factors can be given any meaningful economic interpretation.

The paper concludes that return generating models based on market and industry indexes predict security returns better than the factor models. Furthermore, industry models produce even better forecasts than the market model, indicating that multi-index models better describe security returns than single-index models. Finally, our results suggest that there are substantial differences in the predictive power of the return generating models over time and, consequently, the stationarity of these models over time.

Our paper is organized as follows: Section II describes the empirical design. Section III compares the predictive power of return generating models as estimated by factor analysis with that of the simple market model. Section IV investigates the return generating models based on industry models and Section V concludes the paper.

## I. The Design

The analysis in this paper for the most part follows a two-step procedure. The first step is to develop a return-generating model using one sample of data. The second step is to use this model to predict returns in a subsequent period. A comparison of these predictions with the actual subsequent returns permits an assessment of the relative predictive accuracy of different models. The predictive models used in the first part of this paper are empirically estimated factor models and a standard version of the market model.

To place factor models in perspective, let us begin with the usual market model

$$R_{it} = \alpha_i + \beta_i R_{mt} + \varepsilon_{it} \quad , \tag{1}$$

where $R_{it}$ is the return on asset $i$ during interval $t$, $R_{mt}$ is the return on the market portfolio, $\varepsilon_{it}$ is a mean-zero independent disturbance, and $\alpha_i$ and $\beta_i$ are parameters appropriate to asset $i$. For future reference, note that (1) can be rewritten as

$$R_{it} - E(R_{it}) = \beta_i \big(R_{mt} - E(R_{mt})\big) + \varepsilon_{it} \quad . \tag{2}$$

As usually represented, it is not easy to compare factor models directly with the market model; but with some rewriting of factor models, the two models can be made analogous. The usual representation of a factor model is:

$$R_{it} - E(R_{it}) = \sum_{k=1}^{K} \lambda_{ik} f_{kt} + \eta_{it} \quad , \tag{3}$$

where $K$ is the number of factors, $\lambda_{ik}$ is the so-called factor loading of asset $i$ on factor $k$, $f_{kt}$ is the score or value of factor $k$ during interval $t$, and $\eta_{it}$ is a mean-zero independent disturbance. The expected value of $f_{kt}$ is zero, and it is scaled such that $\sigma(f_{kt})$ is 1.0. In addition, estimation of the factors requires some assumption about the covariances between the different factors, and the assumption that is usually made and will also be made in this paper is that $\text{Cov}(f_{kt}, f_{jt}) = 0,\ k \neq j$ .[4]

In the usual estimation procedures, the factor scores or values are linear sums of the independent variables in (3). Thus,

$$f_{kt} = \sum_{j=1}^{N} \omega_{jk}\left(R_{jt} - E(R_{jt})\right) \quad , \tag{4}$$

where the $\omega_{jk}$'s are called the factor scoring coefficients and $N$ is the number of assets in the analysis.[5] Combining (4) and (3) yields

$$R_{it} - E(R_{it}) = \sum_{k=1}^{K} \lambda_{ik} \sum_{j=1}^{N} \left[\omega_{jk}\left(R_{jt} - E(R_{jt})\right)\right] + \eta_{it} \quad . \tag{5}$$

The weights on the returns, $\omega_{jk}$, can be given portfolio interpretations by rescaling their sum to 1.0. If $W_k$ is defined as $\sum_{j=1}^{N} \omega_{jk}$ and assuming that $W_k \neq 0$, (5) can be rewritten as

$$\begin{aligned} R_{it} - E(R_{it}) &= \sum_{k=1}^{K} \lambda_{ik} W_k \sum_{j=1}^{N} \frac{\omega_{jk}}{W_k} \left[R_{jt} - E(R_{jt})\right] + \eta_{it} \\ &= \sum_{k=1}^{K} \gamma_{ik}\left[I_{kt} - E(I_{kt})\right] + \eta_{it} \quad , \end{aligned} \tag{6}$$

where $\gamma_{ik}$ is the scaled factor loading $\lambda_{ik}W_k$, and $I_{kt}$ is the return index corresponding to factor $k$. This rescaling is equivalent to changing the assumption about the standard deviation of each factor.

If $k = 1$, the factor model can be compared directly to the market model as given by (2). The dependent variables are the same. If the weight multiplying $(R_{jt} - E(r_{jt}))$ in (6) were the market weight for asset $j$ at time $(t-1)$, $I_{it}$ would be the same as $R_{mt}$ and $\gamma_{i1}$ the same as $\beta_i$. More generally, with an appropriate rescaling, each of the factors in any estimated factor model can be interpreted as the return on a specific portfolio, or equivalently as some index of stock price movements.[6] However, as will be demonstrated below, the requirement that the estimated factors be orthogonal forces the weights, on occasion, to be quite different from those associated with usual indexes and those of industry indexes, and thus may make it very difficult to assign meaningful economic interpretations to the factors.[7]

## II. The Initial Evidence

The empirical results in this section utilize monthly returns for twenty portfolios of NYSE stocks from the CRSP tapes. The assignment of firms to portfolios follows the approach used in many prior studies and is very similar to that of Fama and MacBeth [13]. All firms with at least 50 months of complete data in five calendar years and complete data over the next 6 calendar years are ranked according to their beta coefficients as estimated over the first five years of the eleven-year period and then partitioned into 20 portfolios of as close to an equal number of securities each as possible. The number of securities per portfolio starts at 19 in 1936, gradually increases to 35 by 1948, and then remains relatively constant through 1981. The second five years of each eleven-year period are used to estimate the various predictive models, and a comparison of the predictions for each month in the eleventh year with the actual 12 monthly returns provides an evaluation of the relative accuracy of each model. The market index used in estimating the beta coefficients is the value-weighted index of the NYSE provided by CRSP.

The first set of predictions, corresponding to the twelve months in 1936, is based upon various predictive models that have been estimated from the preceding ten years of monthly data from 1926-1935. The second set, corresponding to 1937, is based upon models that have been estimated from the monthly data from 1927-1936. Repeating this process year by year produces monthly predictions for the forty-six years from 1936 through 1981.

Since the work of Brennan [2] indicates the possible presence of at least two factors in portfolios ranked by betas, forming portfolios by ranked beta values should bias the empirical results towards discovering multiple factors. Nonetheless, there is always the possibility that the empirical results might be quite sensitive to the number of securities in the portfolios, the way in which portfolios are formed, and whether portfolios are formed at all. However, replications of most of the following analyses on various samples including individual securities indicate that the general nature of the conclusions to be drawn from this sample of 20 portfolios ranked by betas is quite robust. The last part of this section contains summaries of these replications.

A. Factor Model Predictions

Using the procedure of maximum likelihood, we factor analyze the covariance matrix for each set of portfolios and allow the number of factors to vary from one to five. The covariance matrices themselves are estimated from the monthly data of the second five years of each eleven-year period. For subsequent use, the resulting factor loadings and scores are transformed into portfolio numbers as defined in (6).

On the basis of the Barlett's chi-squared criterion, in no year is a one-factor model adequate and in only five years would more than five factors be required. The median number of factors is three. The frequency distributions of probabilities associated with the chi-squared values are shown in Table 1 and are summarized in panel A of Table 2. Of further note, a detailed analysis of the data underlying these two tables discloses wide swings from year to year in the number of common factors needed to describe the return generating function, even though each pair of adjacent years contains 48 common observations out of 60.

Equation (6) provides the basis for making predictions, or more precisely conditional expected returns, for each of the twelve months in the prediction period for each portfolio. The conditioning variables in (6) are: the indexes or factors, $I_{kt}$; the expected values of these indexes, $E(I_{kt})$; the expected returns on the portfolios, $E(R_{it})$; and the scaled factor loadings, $\gamma_{ik}$. With the exception of the indexes themselves, the values of all of these conditioning variables are taken to be their values as estimated in the second five years of each eleven-year period. The indexes themselves are calculated as the weighted averages of the twenty actual portfolio returns in month t, where the weights are those that were estimated in the second five years of each eleven-year period. As a result, only the index values themselves utilize data from the predicting period.[8]

The mean squared errors of these conditional predictions for each of the twenty portfolios, from 1936 to 1981, provide a measure of the predictive accuracy of the models (Table 3). A five-factor model produces the minimum mean squared error for 16 out of the 20 portfolios. The grand means over all the portfolios also indicate that the conditional predictions of the five-factor model are more accurate than those using a lesser number of factors. The differences between any of the grand means are highly significant (Table 4). Since the number of factors varies only from one to five, it is also possible that a six-factor model might have yielded even lesser mean squared errors. For reasons discussed below, we did not pursue this possibility.[9]

An analysis of the mean squared errors for each portfolio for each year discloses that of the 920 such observations, the minimum mean squared errors are associated with five or more factors 30.54 percent of the time and with four factors 22.28 percent of the time (Panel B of Table 2), suggesting the presence of more factors than revealed by the factor analysis, at least at the five percent level. Interestingly, 11.08 percent of the time, one factor minimizes the mean squared error, whereas the factor analysis itself always finds two or more factors.

To summarize so far, factor models with five or more factors produce on average the minimum mean squared errors. In contrast, the usual chi-squared test for the adequacy of the number of factors points to a median number of factors of three. Further, a detailed examination of each portfolio for each year reveals very little relationship between the number of significant factors in the estimation period and the numbers indicated by the conditional forecasts from one year to the next.[10]

Before accepting at face value the conclusion that the conditional predictions are on average more accurate as the number of factors increases, let us pause to analyze the conditioning process. The calculation of the indexes utilizes the returns of all of the portfolios, including the return of the portfolio for which the prediction is being made. Thus the predictions of the return of a specific portfolio is based in part upon the actual return that is being predicted.

To explore the effect of conditioning a prediction in part upon the return being predicted, we define two instrumental variables. The first is obtained by regressing the return of the asset being predicted on the returns of the 19 other assets during the estimation period. This regression is then used to obtain predictions of the asset's return in the prediction period conditional on the realized returns of the other 19 assets. The second is merely the average of the returns of the other 19 assets in the prediction period. Since the results using this second variable are virtually identical to those using the first, the tables contain only the results for the first method.

Replacing the return being forecasted with its instrumental return in the calculation of the indexes, we replicated the previous tests. The results are quite different from before (columns 7 through 11 on the right hand side of Table 5). Seven of the twenty portfolios have minimum mean squared errors with predictions based upon one or two-factor models. The average mean squared errors over all 20 portfolios is minimized with a three-factor model. There is a statistically significant improvement in the accuracy of the predictions as one moves from a one-factor to a two-factor model and from a two-factor model to a three-factor model (Table 6). Moving from a three- to a four- or a four- to a five-factor model produces little statistical gain. The detailed results by year and by portfolio also point to a fewer number of factors than discovered by factor analysis (Panel C of Table 2).

B. A More Detailed Analysis

To gain some understanding of what happens as the number of factors increases, consider the scaled factor loadings and scoring coefficients for the one-factor and the five-factor models as estimated over the period 1935-1939, an arbitrarily chosen period (Table 7). The factor scoring coefficients are scaled as in (6) to sum to 100 percent. For the one-factor model, the portfolio weights range from 0.6 percent to 11.4 percent, and one's general impression, with the possible exception of the portfolios formed from the lower-ranked betas, is that the weights are fairly equal.

But the story is quite different for the five-factor model. For the first factor, portfolio 12 receives a weight of 82.4 percent with all the other portfolios having weights of between 0.3 and 2.0 percent. For the second factor, portfolio 12 receives a weight of -1992.9 percent, whereas ten of the other portfolios receive weights of over 100 percent. It is almost as if the second factor is undoing the heavy weight given to portfolio 12 in the first factor. The weights in the third through fifth factors are even more extreme.

Since most economic variables are correlated, it is unlikely that factors identified using the assumption of orthogonality will be directly associated with any obvious economic variables. However, if the factors were rotated or estimated under some alternative assumption as to the covariance structure of the factors, the factors might be directly related to some economic variables. But, to make a sensible assumption as to the covariance structure of the factors would seem to require some prespecification of the economic variables. As a result, factor analysis by itself is unlikely to discover those economic variables that are important in explaining returns.

As the number of factors increases to 20, the same as the number of portfolios, it can be shown that a factor model with no residual variance will be able to explain 100 percent of the total variance.[11] Intuitively, one might set each of the factors equal to the return on one of the twenty portfolios and set the response coefficients for each portfolio to one for the factor defined as its own return and zero otherwise. Let us call this model the identity model. This identity model is not a factor model in the usual sense since the factors would generally be correlated.

Since both the identity model and an orthogonal factor model with the same number of factors as portfolios can explain 100 percent of the total variance with no residual variance, it has to be the case that the two models are identical in some sense. Specifically, by rotating the orientation of the factors, one can transform a true factor model into an identity model.[12] Since the factor model with the number of factors equal to the number of portfolios can be transformed into the identity model, any conditional predictions using the actual returns in month t, even outside the estimation period, will be exact.

From the above, it might seem that as the number of factors increases from one and approaches the number of portfolios, the factor models should explain a greater proportion of the total variance. This statement is certainly true in the estimation period since each additional factor is selected mathematically so as to increase the explained proportion of total variance.

However, the accuracy of conditional predictions outside the estimation may not improve with increases in the number of factors as long as the number of factors is less than the number of portfolios. First, if the stochastic process generating returns is non-stationary, there is no guarantee that outside the estimation period, a factor model with more factors would necessarily predict returns more accurately. Second, even if the stochastic process is stationary, there is no mathematical guarantee, only an expectation, that the accuracy of the conditional forecasts would increase with increases in the number of factors.

To examine the joint effect of all of the factor scoring and loading coefficients, note that (5) can be rewritten as

$$R_{it} - E(R_{it}) = \rho_i \sum_{j=1}^{N} \frac{\sum_{k=1}^{K} \lambda_{ik}\omega_{jk}}{\rho_i} \left[R_{jt} - E(R_{jt})\right] \quad , \tag{7}$$

where $\rho_i$ is $\sum_{j=1}^{N} \sum_{k=1}^{K} \lambda_{ik}\omega_{jk}$. In this form, it is apparent that factor analysis is equivalent to specifying a unique index for each portfolio and a response coefficient $\rho_i$ for that portfolio. Only for a one-factor model would the index be the same for all portfolios. Intuitively, what equation (7) does is to transform the dependency built in through the individual orthogonal factors into a set of index weights, generally different for each portfolio.

To examine the effect of conditioning a prediction of a return partially upon the return itself, we determine the weights for the index associated with each portfolio as given in (7) for each of the five different factor models and for each year. Of principal concern is the portfolio weight on the return of the security that is being predicted. The average weight over all portfolios and all years assigned to the return of the asset being forecasted is 0.05 for a one-factor model, 0.096 for a two-factor model, 0.136 for a three-factor model, 0.177 for a four-factor model, and 0.220 for a five-factor model. Thus, as the number of factors increase at least through five, the weight given to the return of the security being predicted increases on average. This is an average effect and does not hold for every portfolio.

C. The Market Model

An alternative model to the factor models is the simple market model as given by (1). To use the market model, one must specify the market index. Two obvious market indexes are the equally-weighted and value-weighted index of NYSE stocks that CRSP calculates. Many prior empirical studies have used one or both of these two indexes. Conditional predictions using these two indexes are subject to the same type of conditioning bias that has been discussed above in that the indexes are based in part on the return of the stock or portfolio whose return is being predicted. For individual stocks, this bias is undoubtedly small except possibly for the larger NYSE stocks in the case of the value-weighted index.

For the portfolios used in this study, the bias may be more important. To eliminate this potential bias, we also construct conditional predictions based upon an equally-weighted index of the 20 portfolios that formed the basis for the prior factor analysis. For each of the twelve predicted months, the return of the portfolio being predicted is replaced in the index by the first variant of its instrumental return as defined above.

The returns from the second five years of each eleven-year period provide the data that are used to estimate the three variants of the market model. For the estimation period, the index of the 20 portfolios is based upon the actual realized returns; instrumental returns are only used in the prediction period. In addition, we used a naive prediction model to provide a reference point, namely the average return of the portfolio in the second five years.

In all cases, the various market models and factor models provide more accurate predictions than the naive model (Table 6). The market model using the equally-weighted NYSE index provides more accurate predictions than using the value-weighted NYSE index. It is interesting to note that the market model using the equally weighted index of the 20 portfolios with the appropriate instrumental returns yields on average forecasts almost as accurate as those based upon the equally-weighted NYSE index. Taken at face value, the market model using the equally-weighted NYSE index provides predictions that are on average more accurate than any of the factor models, but the superiority of this version of the market model is not significant at any usual level of significance except for a one-factor model (Table 7).

In short, the simple market model using an equally-weighted index of NYSE stocks provides as good a prediction as the potentially richer factor models. It remains to be determined how sensitive these results are to the particular way in which portfolios are formed and to the use of portfolios rather than individual securities.

D. Other Experiments

The results reported so far are based upon a specific way of forming portfolios, namely portfolios based upon ranked values of betas. To examine the robustness of these results, we undertake additional experiments that involve some combination of different methods of forming portfolios, the use of individual securities, and different time periods. The general qualitative conclusions from these additional experiments remain unchanged.

The previous portfolios allowed the number of securities to vary over time and the number of portfolios was arbitrarily set to 20. In the following two experiments, the number of securities per portfolio is fixed at 18. The first experiment utilizes 20 portfolios, whereas the second utilizes 30 portfolios.

The design of these two experiments is similar to the previous one. First, all stocks are ranked according to their betas as estimated in the first five years of each 11-year period and divided into 20 or 30 groups of an equal number of stocks. Within each group, 18 stocks are selected at random and without replacement to form a portfolio. Eighteen stocks is the maximum number of stocks for which 30 portfolios could always be formed for every grouping period.

The qualitative results are identical to those reported above (Table 8). Using instrumental returns, three factors produce on average the smallest squared prediction errors. Again, conditional predictions utilizing the equally-weighted NYSE perform very well -- in one experiment better than the factor models and in the other slightly worse.

To determine the sensitivity of the results to the use of portfolios formed by ranked betas, we repeat the previous two experiments except that the 18 stocks for the 20 or 30 portfolios are selected at random and without replacement from the entire universe. Again, the qualitative results, not presented here, are unchanged.

It might be argued that, even if there are multiple factors in explaining the returns of individual securities, a portfolio of securities might hide these multiple factors as some securities are positively affected by a factor and some negatively, offsetting each other in a portfolio. To examine this possibility, we repeat much of the previous analyses but using individual securities rather than portfolios.

Specifically, we randomly select groups of 30 stocks without replacement for every non-overlapping five-year period, starting with 1935. There are 12 such groups in 1935 and 32 groups in 1980. We use each group of 30 stocks to estimate factor loadings and betas to predict returns for each of the stocks. We then compute mean squared prediction errors for each of the same prediction methods as for the portfolios.

For each of the prediction years, we compute the average mean squared error for each of the groups. The grand mean of these average yearly mean squared errors indicates somewhat fewer factors than discovered with portfolios. Of perhaps most interest, the predictions using the equally-weighted NYSE index are more accurate than the predictions based upon estimated factor models. As would be expected, the mean squared prediction errors are larger for individual securities than for portfolios.

The above experiments make conditional predictions outside the estimation period. If the covariance maxtrix changes over time, there is always the possibility that the factor models might predict more accurately within the estimation period than outside the estimation period. To explore this possibility, we estimate the various models utilizing the "odd" months within a period of time and evaluate the models utilizing the "even" months within the same period. This is similar to the approach used by Roll and Ross [21].

For the ten-year period beginning in 1931, we randomly form without replacement 12 groups of 30 individual stocks. For each group, we estimate the various models utilizing the "odd" months and evaluate the models utilizing the "even" months. Previous researchers have reported anomalies in January. In order to minimize any bias from classifying January as an "odd" month continuously, we randomly classified January for each year as an "odd" or an "even" month. If January in any year is used to estimate the models, then March, May and so on of that year would also be used to estimate the models. If January is used to evaluate the models, then March, May and so on would be used to evaluate the models.

We repeat this process for each of the next five non-overlapping periods of ten years ending in 1980. The total number of groups for the last ten years is 24. The grand mean squared prediction errors again lead to the same qualitative conclusions.

An examination by month of the squared prediction errors reveals systematically larger squared errors in January and November. Recomputing the grand mean squared errors without these two months' results, of course, in lower means, but the qualitative conclusions remain unchanged. These results are not reported.

In any analysis of robustness, it is always possible to conceive of other experiments. The experiments presented here do address some of the major issues, such as portfolios versus individual stocks. The general conclusion is that the factor models generally predict most accurately with two or three factors. But predictions using an equally-weighted NYSE index as in the standard one-factor market model do fully as well. Thus, the simple one-factor model is as adequate a model for describing returns as the more complicated factor models.

## III. Industry Models

Despite the apparent inability of factor models to perform significantly better than the simple market model, King [18] and later Meyers [20] have demonstrated the existence of industry effects. It therefore seems appropriate to conclude this paper with an analysis of the predictive power of the market model augmented with industry indexes.

At the end of 1935, there are seven two-digit industries with at least 15 NYSE companies that have a complete monthly return history for the five previous years. To place each industry on the same basis, we randomly selected 15 companies to represent each industry. The two-digit industry codes are those from the CRSP files.

For each company within each industry, several different models are estimated using the data from 1931 through 1935. As with the previous analysis, conditional predictions for the monthly returns in 1936 provide the basis for evaluating the different models. The specific models used are the following:

i) A naive model that uses the average realized return for the prior five years as the prediction.

ii) The usual market model using the value-weighted NYSE index as the measure of the market.

iii) The usual market model using the equally-weighted NYSE index as the measure of the market.

iv) A naive industry model in which the conditional expected return on each security in an industry is the same as the industry average, designated in the tables as "B = 1".

v) An industry model in which the return of a specific security was regressed on an industry index. This is similar to the usual market model except that the industry index replaces the market index.

vi) An augmented market model using the value-weighted NYSE index and an industry index.

vii) An augmented market model using the equally-weighted NYSE index and an industry index.

The index for any industry is defined as the equally-weighted average of the monthly returns of the 15 randomly-selected companies in that industry. As before, the conditional predictions that utilize an industry index based in part upon the actual return of the company whose returns are being predicted produce the best predictions. Thus, to conserve space, only the results using the instrumental return for the security being predicted are presented here.

The instrumental return for any security is obtained by regressing the 60 months of prior returns of that security on the returns of the other 14. The subsequent instrumental return is the expected return of the security conditional on the returns of the other 14 securities.

Table 9 contains the detailed results for the conditional predictions for 1936. Similar results are obtained for every fifth year from 1941 through 1981. A perusal of these tables shows that the mean squared prediction errors for all the models are greatest in 1936, decrease gradually through the mid-fifties, and then increase slightly through 1981 (Table 10). Within an individual industry, there are widely disparate time trends with the mean squared errors of some industries increasing, some decreasing, and some oscillating over time.

Moreover, there are major differences among industries. As might be expected, the mean squared errors for some industries are consistently lower than for other industries. There are seven industries that are represented in each of the ten prediction years. The mean squared errors for these industries averaged over the ten prediction years are shown in Table 11.

Although the mean squared errors display both time and industry effects, the differences between the mean squared errors within any time period and industry for any pair of prediction techniques do not appear to be systematically related to time or industry. Applying standard t-tests to these differences discloses no significant difference between an industry model with a beta of 1.0 for all firms and the industry model based upon a regression of the firm's returns on an industry index (Table 12). The t-values for the differences between any of the other prediction models and either of these industry models are all greater than 2.0.

What this means is that an industry index predicts the returns of a security better than the standard market index. The implication is that returns are better explained by a multi-index model than by a single-index model. Since industry indexes are highly correlated, factor analysis as usually applied with the assumption of orthogonality would not be able to discover these factors. Only by rotating the factors would one be able to discover industry effects, and to do this one would have to have some hypothesis as to the covariance matrix of industry indexes.

Also of interest is that the market model augmented with an industry index performs less well than just using the industry index alone. Apparently, what is happening is that there is enough non-stationarity in the return-generating process that a regression involving both a market and an industry variable only adds some further noise to an already noisy process.

## IV. Conclusion

This paper has demonstrated that the returns of individual securities are better described by a multifactor model where the factors are related to industry classifications. It also has shown that a blind application of factor analysis can lead to less accurate models than the usual market model. Further, the orthogonality assumption of factor analysis makes it very unlikely that the factors discovered by factor analysis will have any economically meaningful interpretation.

## FOOTNOTES

1. There is a new line of research testing arbitrage pricing models in which the researcher specifies a set of observable economic variables as "fundamental factors" underlying the security returns. See, for example, Chen, Roll and Ross [6] or Chan, Chen and Hsieh [4]. The methodology we use in this paper can easily be extended to incorporate any return-generating model including those that are based on a set of observable economic variables. In a recent project, we are indeed examining the properties of return-generating models based on observable economic variables.

2. The previous tests of the APT either usually assume that the number of factors are equal to some predetermined number or attempt to estimate the number simultaneously with the estimation of factor loadings. There are few attempts to determine the number of factors separately from the cross-sectional pricing tests. Such papers include those by Dhrymes, Friend and Gultekin [9], Gibbons [16], and Trizcinka [23].

3. The reason for examining predictions rather than the sampling statistics of the factor analysis itself is that violations of the statistical analysis underlying factor analysis may bias these statistics in unknown ways. For example, maximum likelihood estimates of factor structure assume a stationary normal process independent over time. There is strong evidence, however, that stock returns do not conform to a stationary normal process, and thus there is always the possibility that factor analysis may discover more or less factors than are present.

4. Note that there is a very important inherent difference between the market and factor models. In the market model, $R_{mt}$ is an observable variable and $\beta$ is estimated. In the factor models, neither $\lambda_{ik}$ nor $f_{kt}$ are observable. This creates an identification problem which is partly eliminated by specifying that $E(f_{kt}) = 0$ and $\sigma(f_{kt}) = 1.0$ for all $k$ and $cov(f_{it}, f_{jt}) = 0$ for all $i \neq j$. Since only the covariance matrix is estimated directly from the data, there is a further identification problem. Factor loadings $\lambda_{ik}$ can be identified only up to left multiplication by an orthogonal matrix.

5. For later use, it is convenient to express factor models in matrix notation as follows:

Let

$\mathbf{R}_t' = [R_{1t}, R_{2t}, \ldots, R_{nt}]$ n-element vector of returns ;

$\mathbf{f}_t' = [f_{1t}, f_{2t}, \ldots, f_{kt}]$ k-element vector of factor scores ;

$\eta_t' = [\eta_{1t}, \eta_{2t}, \ldots, \eta_{nt}]$ n-vector of specific-factor variate ;

and $\mathbf{\Lambda}$ is a $n \times k$ matrix containing factor loadings.

Now equation (2) can be rewritten as:

$$\mathbf{R}_t - E(\mathbf{R}_t) = \mathbf{\Lambda f}_t + \eta_t \quad ,$$

with

$$\text{cov}(\eta_t) = \mathbf{\Psi} \quad ,$$

a diagonal $n \times n$ matrix, and

$$\text{cov}(\mathbf{R}_t - E(\mathbf{R}_t)(\mathbf{R}_t - E(\mathbf{R}_t)') = \mathbf{\Lambda\Lambda}' + \mathbf{\Psi} \quad .$$

Factor scores, $\mathbf{f}_t$, are usually estimated as

$$\hat{\mathbf{f}}_t = \hat{\mathbf{\Lambda}}'(\hat{\mathbf{\Lambda}}\hat{\mathbf{\Lambda}}' + \hat{\mathbf{\Psi}})^{-1}(\mathbf{R}_t - E(\mathbf{R}_t)) = \mathbf{S}^{-1}\hat{\mathbf{\Lambda}}(\mathbf{R}_t - E(\mathbf{R}_t)) \quad ,$$

where the hat above a variable indicates the maximum likelihood of estimates of the true parameters and $\mathbf{S}$ is the sample covariance matrix of returns.

Note that weight $\omega_{jk}$ in equation (4) is an element of $\mathbf{S}^{-1}\hat{\mathbf{\Lambda}}$ matrix.

6. Since factor analysis assumes that the weights on the individual returns are constant over time, the estimated factors would never correspond exactly to an index like the Standard & Poor's Composite Index of 500 stocks where the weights change from one month the the next. Nonetheless, King's work [18] suggests that this theoretical point may not be empirically important.

7. The assumption of orthogonality arises from the identification problem. Note that in factor analysis, both factor loadings, $\lambda_{ik}$'s , and factor scores, $f_t$'s, are unobservable and thus simultaneously estimated. The assumption about the covariance structure among factor scores eliminates one part of the identification problem. (See Footnote 1.) There are procedures, however, which do not require that factor scores be orthogonal.

FOOTNOTES (cont.)

8. Predictions and mean-squared forecast errors for a given year $t$ are computed as follows. First, factor scores are estimated for the prediction period by

$$\hat{\mathbf{f}}_{t+s} = \hat{\Lambda}'(\hat{\Lambda}\hat{\Lambda}' + \hat{\Psi})^{-1}\left(R_{t+s} - E(R_t)\right) \qquad s = 1,2,\ldots,12 \quad .$$

$\hat{\Lambda}$ , matrix of factor loadings, and $\hat{\Psi}$ , residual variances, are estimated using the five years of observations prior to the prediction period. $E(\mathbf{R})$ is the vector of mean returns and is estimated by the arithmetic average of returns using the five years of returns prior to the prediction period. $\mathbf{R}_{t+s}$ is the observed market return during the prediction period. The next stage is to predict the security returns using the estimated factor scores $\hat{\mathbf{f}}_{t+s}$ , i.e.

$$\hat{R}_{t+s} - E(R_t) = \hat{\Lambda}\hat{\mathbf{f}}_{t+s} \quad .$$

Mean-squared prediction errors then are computed as the difference between the actual and predictions:

$$\text{MSPE} = (1/M)(\hat{\mathbf{R}}_{t+s} - \mathbf{R}_{t+s})'(\hat{\mathbf{R}}_{t+s} - \mathbf{R}_{t+s}) \quad ,$$

where $M$ is the size of vector $\mathbf{R}_t$ containing stock returns.

9. Also of interest is that the mean squared errors for those portfolios formed of securities with the extreme values of beta in the grouping period, either large or small, tend to be greater than for those portfolios with less extreme betas. One possible explanation is that the return distributions of those securities with extreme betas are less stationary than those with less extreme betas, an explanation that is consistent with the regression tendencies of betas as evidenced in Blume [1].

10. For example, if one uses the sufficient number of factors in a given five-year period as a forecast for the next five years after updating the sample by one year, only 16 times out of 46 years would this forecast would have been right.

11. Factor analysis turns into principal component analysis in this special case. The principal component analysis can be considered a special case of factor analysis where specific variances are all assumed to be zero.

12. This statement assumes that variance-covariance matrix has full rank. If the number of factors equals the number of returns, from Footnote 5, we have

$$R_t - E\left(R_t\right) = \left(\hat{\Lambda}\hat{\Lambda}'\right)\left(\hat{\Lambda}\hat{\Lambda}'\right)^{-1}\left(R_t - E(R_t)\right) = I\left(R_t - E(R_t)\right) \quad ,$$

where $I$ is the identity matrix.

## REFERENCES


1. Marshall E. Blume. "On the Assessment of Risk." Journal of Finance 26 (March 1971), 1-10.

2. Michael J. Brennan. "Capital Asset Pricing and the Structure of Security Returns." Unpublished Manuscript, University of British Columbia, May 1971.

3. Stephen J. Brown and Mark Weinstein. "A New Approach to Testing Asset Pricing Theories: The Bilinear Paradigm." Journal of Finance 38 (June 1983), 711-43.

4. K.C. Chan, Nai-fu Chen, and David A. Hsieh. "An Exploratory Investigation of the Firm Size Effect," Journal of Financial Economics 14 (September 1985), 451-71.

5. Nai-fu Chen. "Some Empirical Tests of the Theory of Arbitrage Pricing." Journal of Finance 38 (December 1983), 1392-1414.

6. Nai-fu Chen, Richard Roll, and Stephen A. Ross. "Economic Forces and the Stock Market: Testing the APT and Alternative Asset Pricing Models." CRSP Working Paper 119, University of Chicago, 1984.

7. David C. Cho, Edwin J. Elton, and Martin J. Gruber. "On the Robustness of the Roll and Ross APT Methodology." Journal of Financial and Quantitative Analysis 19 (March 1984), 1-10.

8. Gregory Conner. "A Unified Beta Pricing Theory." Journal of Economic Theory 34 (December 1984), 13-31.

9. Phoebus J. Dhrymes, Irwin Friend and N. Bulent Gultekin. "A Critical Reexamination of the Empirical Evidence on the Arbitrage Pricing Theory." Journal of Finance 39 (June 1984), 323-46.

10. Phoebus J. Dhrymes, Irwin Friend, N. Bulent Gultekin and Mustafa N. Gultekin. "An Empirical Examination of the Implications of Arbitrage Theory." Journal of Banking and Finance 9 (March 1985), 73-99.

11. Edwin J. Elton and Martin J. Gruber. "Improved Forecasting Through the Design of Homogeneous Groups." Journal of Business 44 (October 1971), 432-50.

12. Edwin J. Elton and Martin J. Gruber. "Estimating the Dependence Structure of Share Prices - Implications for Portfolio Selection." Journal of Finance 8 (December 1973), 1203-32.

13. Eugene F. Fama and James D. MacBeth. "Risk, Return and Equilibrium: Empirical Tests." Journal of Political Economy 38 (May 1973), 607-36.

14. Donald E. Farrar. The Investment Decision Under Uncertainty, Englewood Cliffs, NJ: Prentice-Hall, Inc., 1962.

15. James Farrell. "Analyzing Covariation of Returns to Determine Homogeneous Stock Groupings." Journal of Business 47 (April 1974), 186-207.

## REFERENCES (continued)

16. Michael R. Gibbons. "Empirical Examination of the Return-Generating Process of the Arbitrage Pricing Theory." Research Paper Series, Graduate School of Business, Stanford University, May 1986.

17. N. Bulent Gultekin and Richard J. Rogalski. "Government Bond Returns, Measurement of Interest Rate Risk, and the Arbitrage Pricing Theory." Journal of Finance 40 (March 1985), 43-61.

18. Benjamin F. King. "Market and Industry Factors in Stock Price Behavior." Journal of Business 39 (January 1966), 139-90.

19. Miles Livingston. "Industry Movements of Common Stocks." Journal of Finance 32 (June 1977), 861-74.

20. Stephen Meyers. "A Re-examination of Market and Industry Factors in Stock Price Behavior." Journal of Finance 8 (June 1973), 695-709.

21. Richard Roll and Stephen A. Ross. "An Empirical Investigation of the Arbitrage Pricing Theory." Journal of Finance 35 (December 1980), 1073-1103.

22. Stephen A. Ross. "The Arbitrage Theory of Capital Asset Pricing." Journal of Economic Theory 13 (December 1976), 341-60.

23. Charles Trzcinka. "On the Number of Factors in the Arbitrage Pricing Model." Journal of Finance 41 (June 1986), 347-68.

**TABLE I**

CROSS SECTIONAL FREQUENCY DISTRIBUTION OF PROBABILITIES ASSOCIATED WITH CHI-SQUARED STATISTICS FROM A LIKELIHOOD RATIO TEST THAT NO MORE THAN k FACTORS ARE NECESSARY TO EXPLAIN THE RETURNS ON 20 PORTFOLIOS OF NYSE STOCKS[a]

1936-1981

| # of factors:[b] k | Probability that no more than k factors are needed[c] 0-.05 | .05-.1 | .1 | .2 | .3 | .4 | .5 | .6 | .7 | .8 | .9 | Number of Times Exactly k Factors are Sufficient[d] |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 1 | 46 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 2 | 38 | 4 | 1 | 3 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 8 |
| 3 | 21 | 7 | 7 | 3 | 1 | 1 | 4 | 0 | 0 | 1 | 0 | 17 |
| 4 | 11 | 1 | 10 | 3 | 8 | 4 | 1 | 3 | 2 | 2 | 1 | 10 |
| 5 | 5 | 1 | 5 | 5 | 4 | 3 | 5 | 7 | 5 | 1 | 6 | 6 |

a Portfolios are formed from securities that are ranked with respect to their betas. Portfolio 1 contains the securities with the lowest betas and Portfolio 20 contains the securities with the highest betas. Each portfolio contains 5 percent of NYSE stocks in any year.

b Factors are estimated for a given year using the observations from the previous five years. Estimates are updated each year. Each factor model is estimated separately.

c The probability value indicates that the test statistic (under the null hypothesis) will assume a value at least as large as the statistic obtained in this particular test. The null hypothesis that k factors are sufficient is accepted if the p value is greater than .05. For one factor, for example, the hypothesis is rejected every year.

d Indicates the number of years that exactly a k-factor model is sufficient to describe the portfolio returns.

## TABLE II

### PERCENT OF YEARS FOR WHICH A k-FACTOR MODEL IS SUFFICIENT AND THE FACTOR MODELS WITH MINIMUM MEAN SQUARED FORECAST ERROR (MSFE)

| | Number of Factors | | | | |
|---|---|---|---|---|---|
| | 1 | 2 | 3 | 4 | 5 (or more) |
| A. Percent of years for which k-factor model is sufficient[a] | | | | | |
| | 0.0 | 17.39 | 36.95 | 21.73 | 23.91 |
| B. Percent of years for which k-factor produces the smallest MSFE[b] (actual returns) | | | | | |
| | 11.08 | 17.60 | 18.48 | 22.28 | 30.54 |
| C. Percent of years for which k-factor produces the smallest MSFE (instrumental returns) | | | | | |
| | 15.10 | 25.97 | 20.21 | 18.01 | 20.65 |

[a] Figures indicate the percent of years for which no more than k-factors are needed to explain the 20 common stock portfolios returns.

[b] Mean squared forecast error is computed for each portfolio every year. The number of factors that is associated with the smallest MSFE for each portfolio constitutes an observation. There are 920 such observations (46 years for 20 portfolios a year).

TABLE III

COMPARISON OF MEAN SQUARED PREDICTION ERRORS (MSPE) FOR FACTOR MODELS[a]
1/1936-12/1981

| Portfolio[b] Rank (1) | Factor Models[c,d] k = 1 (2) | k = 2 (3) | k = 3 (4) | k = 4 (5) | k = 5 (6) |
|---|---|---|---|---|---|
| 1 | 11.18 | 7.18 | 6.62 | 6.98 | 4.03 |
| 2 | 7.37 | 3.34 | 2.30 | 2.13 | 2.03 |
| 3 | 7.23 | 2.40 | 1.95 | 1.85 | 1.59 |
| 4 | 7.46 | 2.26 | 1.74 | 1.81 | 1.69 |
| 5 | 6.10 | 2.24 | 1.91 | 1.84 | 1.76 |
| 6 | 4.75 | 2.09 | 2.09 | 2.45 | 2.71 |
| 7 | 3.44 | 2.27 | 2.02 | 2.05 | 1.80 |
| 8 | 3.33 | 2.34 | 2.29 | 2.26 | 2.22 |
| 9 | 5.98 | 6.04 | 5.02 | 3.69 | 3.03 |
| 10 | 2.63 | 2.54 | 2.35 | 2.25 | 2.36 |
| 11 | 2.55 | 2.61 | 2.55 | 2.71 | 2.54 |
| 12 | 1.89 | 2.13 | 2.08 | 1.91 | 1.52 |
| 13 | 2.49 | 2.19 | 2.03 | 1.89 | 1.95 |
| 14 | 2.66 | 2.87 | 2.75 | 2.41 | 2.39 |
| 15 | 2.38 | 2.07 | 1.96 | 1.96 | 1.87 |
| 16 | 3.67 | 2.88 | 2.80 | 2.72 | 2.36 |
| 17 | 6.50 | 3.49 | 3.43 | 2.93 | 2.79 |
| 18 | 5.90 | 3.44 | 3.20 | 2.88 | 2.83 |
| 19 | 9.85 | 6.62 | 6.12 | 3.77 | 3.43 |
| 20 | 8.92 | 6.12 | 4.14 | 3.91 | 3.67 |
| Grand MSFE | 5.31 | 3.36 | 2.97 | 2.72 | 2.43 |
| Sqr. GMSFE | 2.31 | 1.83 | 1.72 | 1.65 | 1.56 |

a All returns are measured as percentages. Grand MSPE is the average MSPE of each column. Sqr. GMSPE is the square root of Grand MSPE.

b Portfolio 1 contains stocks with the lowest betas in first 5th percentile and portfolio 20 with the largest betas in the last 5th percentile of the sample. Each portfolio contains 5 percent of NYSE stocks in any year.

c Factor model forecasts for each month $s$ of the year $t$ is obtained by $\hat{\mathbf{R}}_{t+s} = E(\mathbf{R}_t) + \hat{\mathbf{\Lambda}}\hat{\mathbf{f}}_{t+s}$, where $\hat{R}_{t+s}$ is a 20-element column vector containing returns predictions. $\mathbf{\Lambda}$ is a 20 x k matrix of factor loadings estimated using the monthly observations from t-59 to t. $\hat{\mathbf{f}}_{t+s}$ is a k-element column vector of predicted factor scores for each month $s$ of the year $t$ and $s = 1,2,\ldots,12$. $\hat{\mathbf{f}}_{t+s}$ is estimated by $\hat{\mathbf{f}}_{t+s} = \hat{\mathbf{\Lambda}}'(\hat{\mathbf{\Lambda}}\hat{\mathbf{\Lambda}}' + \hat{\mathbf{\Psi}})\mathbf{r}_{t+s}$, where $r$ represents deviation from the mean such that $\mathbf{r}_{t+s} = R_{t+s} - E(\mathbf{R}_t)$. $E(\mathbf{R}_t)$ is estimated by the average of five year returns (from t-59 to t) prior to the forecast period.

d Factor loadings are estimated separately for each k.

## TABLE IV

SIGNIFICANCE TESTS FOR THE DIFFERENCES IN MSPE's AMONG 1-5 FACTOR MODELS: t-RATIOS[a]

| Factor Models[b] (1) | Factor Models k = 2 (2) | k = 3 (3) | k = 4 (4) | k = 5 (5) |
|---|---|---|---|---|
| k = 1 | 3.72 | 4.09 | 4.05 | 4.13 |
| k = 2 | | 5.46 | 4.83 | 4.32 |
| k = 3 | | | 3.45 | 3.54 |
| k = 4 | | | | 3.02 |

a t-ratios are computed to test the significance of the differences among the Grand MSPEs in Table 3.

b The differences in MSPEs are measured between the model shown in each row of column (1) and the 2 to 5 factor models in columns (2) through (5). t-ratio is computed by $(\bar{d}/\sigma_d)\sqrt{T}$, where $d = (1/T)(SPE_t^{(i)} - SPE_t^{(j)})$ and $\sigma_d = (1/T)\sum(d_t - \bar{d})^2$. $SPE_t^{(i)}$ is the squared prediction error for the i-factor model at period t.

## TABLE V

COMPARISON OF MEAN SQUARED PREDICTION ERRORS BETWEEN FACTOR MODELS AND VARIOUS MARKET MODEL FORECASTS:
REPLACING OWN RETURNS WITH INSTRUMENTAL RETURNS
1936-1981

| | | | Market Models[e] | | | Factor Models[f] | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| Portfolio Rank[b] (1) | Naive[c] (2) | b = 1[d] (3) | VW (4) | EW (5) | IR (6) | k = 1 (7) | k = 2 (8) | k = 3 (9) | k = 4 (10) | k = 5 (11) |
| 1 | 16.26 | 12.58 | 5.96 | 4.58 | 4.68 | 11.26 | 7.48 | 7.24 | 7.80 | 5.38 |
| 2 | 20.58 | 6.74 | 4.55 | 3.07 | 3.06 | 7.42 | 3.74 | 2.90 | 2.77 | 2.94 |
| 3 | 22.17 | 6.34 | 4.24 | 3.13 | 3.12 | 7.26 | 2.87 | 2.56 | 2.72 | 2.53 |
| 4 | 22.71 | 6.41 | 3.47 | 3.42 | 3.44 | 7.61 | 2.71 | 2.49 | 2.61 | 2.50 |
| 5 | 25.23 | 4.93 | 4.33 | 2.96 | 2.90 | 6.17 | 2.53 | 2.57 | 2.77 | 2.90 |
| 6 | 28.14 | 3.83 | 4.06 | 2.58 | 2.57 | 4.93 | 2.43 | 2.63 | 3.25 | 3.67 |
| 7 | 29.72 | 2.67 | 4.60 | 2.06 | 1.96 | 3.59 | 2.49 | 2.29 | 2.37 | 2.33 |
| 8 | 33.31 | 2.80 | 4.56 | 2.69 | 2.47 | 3.65 | 2.88 | 3.05 | 3.08 | 3.09 |
| 9 | 46.91 | 7.23 | 13.42 | 7.02 | 6.93 | 6.66 | 6.65 | 6.26 | 5.99 | 6.17 |
| 10 | 39.47 | 2.75 | 6.83 | 2.85 | 2.73 | 2.90 | 2.78 | 2.66 | 2.91 | 3.15 |
| 11 | 45.13 | 3.11 | 8.36 | 3.03 | 2.86 | 2.87 | 2.97 | 3.01 | 3.52 | 3.61 |
| 12 | 44.79 | 2.36 | 6.46 | 2.32 | 2.24 | 2.10 | 2.33 | 2.33 | 2.35 | 2.44 |
| 13 | 40.15 | 2.60 | 7.38 | 2.58 | 2.94 | 2.83 | 2.50 | 2.52 | 2.84 | 3.07 |
| 14 | 48.65 | 3.49 | 8.47 | 2.92 | 3.22 | 3.09 | 3.44 | 3.38 | 3.37 | 3.48 |
| 15 | 50.68 | 3.27 | 8.34 | 2.62 | 2.92 | 2.68 | 3.40 | 2.54 | 2.55 | 2.69 |
| 16 | 57.39 | 4.89 | 11.43 | 3.46 | 3.46 | 3.97 | 3.50 | 3.52 | 3.60 | 3.43 |
| 17 | 67.23 | 8.89 | 15.66 | 5.43 | 5.31 | 7.12 | 4.65 | 4.74 | 4.37 | 4.35 |
| 18 | 68.29 | 7.99 | 15.63 | 4.22 | 4.37 | 6.33 | 4.17 | 4.17 | 4.02 | 4.11 |
| 19 | 77.06 | 11.79 | 17.72 | 6.55 | 6.65 | 10.32 | 7.66 | 7.41 | 7.73 | 7.11 |
| 20 | 73.76 | 10.44 | 18.65 | 6.13 | 6.76 | 9.30 | 7.08 | 6.88 | 6.96 | 7.21 |
| Grand MSPE | 42.87 | 5.76 | 8.71 | 3.68 | 3.73 | 5.60 | 3.86 | 3.76 | 3.88 | 3.81 |
| Sqr. GMSPE | 6.55 | 2.40 | 2.95 | 1.92 | 1.93 | 2.37 | 1.97 | 1.94 | 1.97 | 1.95 |

a All returns are measured as percentages. Grand MSPE is the average of MSPE of columns. Sqr. GMSPE is the square root of GMSPE.

b Portfolios are ranked according to betas in ascending order. Portfolio 1 contains stocks with the lowest betas on the first 5th percentile and portfolio 20 with the largest betas in the last 5th percentile of the sample. Each portfolio contains 5 percent of NYSE stocks in any year.

c Naive forecast is the mean of returns for the previous five years prior to the forecast period.

d $b_i$'s are assumed to be one for each portfolio $i = 1,2,\ldots,20$.

e Market model forecasts for each month $s$ for the year $t$ are obtained by $\dot{r}_{it+s} = \dot{b}_i r_{mt+s}$, $s = 1,2,\ldots,12$. $\dot{b}_i$ is the beta for portfolio $i$ and it is estimated using five years of returns prior to estimation period. $r_m$ is the observed market return. For $r_m$, VW designates the value-weighted index of New York Stock Exchange (NYSE) stocks, EW the equally-weighted index of NYSE stocks, IR is the instrumental variable which consists of equally-weighted index of 20 portfolios where the ith portfolio's return is replaced by its conditional estimate. Returns are deviations from the means, therefore there is no intercept term for the forecast model.

f Factor model forecasts for the year $t$ are obtained by $R_{t+s} = E(R_t) + \dot{G}\dot{f}_{t+s}$. $\dot{R}_{t+s}$ is a 20-element column vector containing return predictions. $\dot{G}$ is a 20xk matrix of factor loadings estimated for the five-year period prior to forecast date. $\dot{f}_{t+s}$ is a k-element column vector of predicted factor scores for the month $s$ in year $t$ and $s = 1,2,\ldots,12$. $\dot{f}_{t+s}$ is estimated by $\dot{f}_{t+s} = \dot{G}'(\dot{G}\dot{G}' + \dot{C})r^*_{t+s}$, where $r^*$ is a vector of portfolio returns at month $s$ of year $t$ measured in deviating form the mean of previous five years and the ith portfolio's return is replaced by its instrumental return. It is estimated by an OLS regression.

## TABLE VI

SIGNIFICANCE TESTS FOR THE DIFFERENCES IN MSPEs BETWEEN FACTOR AND MARKET MODEL FORECASTS: T-RATIOS WHEN OWN RETURNS ARE REPLACED BY INSTRUMENTAL RETURNS[a]

| | | | | | Forecast Models | | | | |
|---|---|---|---|---|---|---|---|---|---|
| | $\beta$ = 1 | VW | EW | IR | k=1 | k=2 | k=3 | k=4 | k=5 |
| (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) | (9) | (10) |
| Naive | 8.01 | 7.66 | 7.74 | 7.73 | 8.01 | 7.82 | 7.79 | 7.81 | 7.74 |
| $\beta$ = 1 | | -4.31 | 3.66 | 3.61 | 2.47 | 3.66 | 3.81 | 3.81 | 8.65 |
| VW | | | -5.24 | 5.19 | 5.15 | 5.06 | 5.15 | 5.14 | 5.07 |
| EW | | | | -1.78 | -3.47 | -1.01 | -.46 | -1.00 | -.83 |
| 20 PS | | | | | -3.41 | -.73 | -.15 | -.77 | -.52 |
| k = 1 | | | | | | 3.47 | 3.58 | 3.50 | 3.43 |
| k = 2 | | | | | | | 1.98 | -.19 | .51 |
| k = 3 | | | | | | | | -1.81 | -.66 |
| k = 4 | | | | | | | | | -.68 |

a t-ratios are for the differences among the Grand MSPEs in Table 6.

b The differences are measured between the models shown in each row of column (1) and the models in columns (2) through (10). t-ratios are computed by $(\bar{d}/\sigma_d)(\sqrt{T})$, where $\bar{d} = (1/T)(SPE^{(i)} - SPE^{(j)})$ and $\sigma_d = (1/T)\sum(d_t-\bar{d})^2$. $SPE^{(i)}$ is the squared prediction error for the ith model at period i.

**TABLE VII**

COMPARISON OF SCALED FACTOR LOADINGS AND SCORING COEFFICIENTS
FOR ONE AND FIVE FACTOR MODELS (1935 - 1939)

| Portfolio Rank | One Factor Model | Five Factor Model | | | | |
|---|---|---|---|---|---|---|
| | | First | Second | Third | Fourth | Fifth |
| **A. Factor Loading Scaled to Portfolio Interpretation** | | | | | | |
| 1 | 0.8526 | 0.8710 | -0.0102 | 0.0453 | -0.0014 | 0.0369 |
| 2 | 0.9371 | 0.9417 | 0.0001 | 0.0431 | -0.0008 | 0.0018 |
| 3 | 0.9587 | 0.9493 | 0.0118 | 0.0344 | 0.0028 | 0.0103 |
| 4 | 0.9234 | 0.9445 | -0.0105 | 0.0360 | 0.0009 | -0.0124 |
| 5 | 0.9651 | 0.9520 | 0.0138 | 0.0313 | 0.0020 | -0.0062 |
| 6 | 0.9702 | 0.9714 | 0.0056 | 0.0234 | -0.0001 | -0.0008 |
| 7 | 0.9911 | 0.9680 | 0.0233 | 0.0169 | 0.0030 | -0.0038 |
| 8 | 1.0056 | 0.9627 | 0.0384 | 0.0074 | 0.0013 | -0.0001 |
| 9 | 1.0015 | 0.9564 | 0.0422 | -0.0019 | 0.0001 | 0.0094 |
| 10 | 1.0064 | 0.9743 | 0.0327 | -0.0009 | 0.0023 | 0.0057 |
| 11 | 0.9952 | 0.9435 | 0.0481 | -0.0072 | 0.0026 | 0.0058 |
| 12 | 0.9823 | 1.0085 | -0.0071 | -0.0011 | 0.0000 | -0.0001 |
| 13 | 1.0042 | 0.9677 | 0.0329 | 0.0188 | 0.0000 | -0.0083 |
| 14 | 0.9946 | 0.9694 | 0.0253 | 0.0161 | -0.0036 | -0.0008 |
| 15 | 1.0103 | 0.9641 | 0.0425 | 0.0007 | -0.0003 | -0.0071 |
| 16 | 0.9882 | 0.9393 | 0.0481 | -0.0190 | 0.0026 | 0.0040 |
| 17 | 1.0064 | 0.9569 | 0.0458 | -0.0041 | 0.0002 | -0.0077 |
| 18 | 1.0068 | 0.9612 | 0.0431 | -0.0031 | -0.0025 | 0.0050 |
| 19 | 1.0087 | 0.9733 | 0.0343 | 0.0002 | -0.0015 | -0.0032 |
| 20 | 0.9934 | 0.9716 | 0.0267 | -0.0063 | -0.0046 | 0.0020 |
| **B. Factor Scoring Coefficients Scaled to Portfolio Weights (Percent)** | | | | | | |
| 1 | 0.6 | 0.3 | -10.1 | 228.5 | -389.0 | 728.6 |
| 2 | 1.3 | 0.5 | 0.2 | 432.5 | -432.2 | 70.7 |
| 3 | 1.8 | 0.5 | 22.7 | 339.2 | 1466.0 | 397.1 |
| 4 | 1.1 | 0.4 | -16.4 | 286.3 | 377.2 | -386.0 |
| 5 | 2.0 | 0.4 | 21.0 | 243.6 | 832.1 | -189.3 |
| 6 | 2.3 | 0.5 | 9.0 | 191.5 | -58.6 | -26.9 |
| 7 | 3.9 | 0.6 | 50.1 | 185.8 | 1757.9 | -165.5 |
| 8 | 7.7 | 0.8 | 104.8 | 103.2 | 974.3 | -6.8 |
| 9 | 6.1 | 0.8 | 120.0 | -27.3 | 91.2 | 532.8 |
| 10 | 8.2 | 1.2 | 133.9 | -18.4 | 2607.8 | 468.4 |
| 11 | 4.5 | 0.9 | 163.0 | -125.1 | 2413.5 | 396.5 |
| 12 | 3.0 | 82.4 | -1992.9 | -1608.7 | 1162.0 | -291.7 |
| 13 | 7.1 | 1.3 | 147.7 | 431.9 | 4.3 | -742.6 |
| 14 | 4.4 | 0.8 | 72.7 | 236.7 | -2825.7 | -43.7 |
| 15 | 11.4 | 1.6 | 246.3 | 20.6 | -416.0 | -824.5 |
| 16 | 3.6 | 1.3 | 234.6 | -473.3 | 3450.4 | 393.9 |
| 17 | 8.2 | 1.6 | 257.0 | -118.8 | 279.2 | -858.8 |
| 18 | 8.5 | 2.0 | 313.9 | -115.0 | -5084.9 | 734.2 |
| 19 | 9.8 | 1.0 | 125.5 | 4.6 | -1525.8 | -232.2 |
| 20 | 4.2 | 1.0 | 96.9 | -117.7 | -4583.7 | 145.8 |

## TABLE VIII

COMPARISON OF MEAN SQUARE PREDICTION ERRORS TO EVALUATE ROBUSTNESS USING INSTRUMENTAL RETURNS

| | | Market Models[a] | | | Factor Models[b] | | | | |
|---|---|---|---|---|---|---|---|---|---|
| Method | Naive | VW | EW | IR | k = 1 | k = 2 | k = 3 | k = 4 | k = 5 |
| A: 20 portfolios of 18 stocks each | 6.57 | 3.18 | 2.28 | 2.31 | 2.65 | 2.34 | 2.31 | 2.33 | 2.36 |
| B: 30 portfolios of 18 stocks each | 5.56 | 2.87 | 2.09 | 2.10 | 2.28 | 2.04 | 2.02 | 2.04 | 2.07 |
| C: Random samples of 30 individual stocks | 9.42 | 8.32 | 8.08 | 8.24 | 8.26 | 8.32 | 8.38 | 8.52 | 8.65 |
| D: Odd-even periods: random samples of 30 individual stocks | 15.82 | 9.55 | 9.21 | 9.34 | 9.61 | 9.55 | 9.65 | 9.72 | 9.80 |

a,b See Table 5 for explanations.

# TABLE IX

## COMPARISON OF INDUSTRY AND MARKET MODELS
## MEAN SQUARE PREDICTION ERRORS FOR INDUSTRY GROUPS
## 1936

| Industry Code[a] | Naive[b] | VW | EW[c] | $\beta$ = 1[d] | II[e] | II & EW[f] | II & VW[g] |
|---|---|---|---|---|---|---|---|
| 20 | 222.26 | 185.29 | 151.03 | 202.05 | 164.49 | 183.44 | 196.05 |
| 28 | 80.01 | 61.21 | 56.30 | 62.31 | 59.47 | 62.38 | 58.66 |
| 29 | 175.06 | 158.58 | 140.67 | 124.86 | 121.97 | 124.94 | 120.66 |
| 33 | 157.61 | 129.95 | 128.19 | 123.95 | 125.13 | 133.80 | 132.63 |
| 35 | 168.67 | 145.10 | 135.13 | 148.84 | 149.95 | 148.68 | 155.48 |
| 37 | 164.38 | 135.29 | 127.16 | 134.36 | 128.77 | 141.56 | 135.85 |
| 40 | 293.06 | 226.69 | 198.90 | 200.38 | 191.08 | 181.55 | 183.10 |
| Average Sum | 1261.06 | 1042.10 | 937.37 | 996.75 | 940.86 | 986.34 | 982.42 |

a Industry code designates the two-digit SIC code.

b Naive forecast is the simple arithemetic mean of returns for the five years prior to forecast period.

c The market indexes are the value-weighted (VW) and equally-weighted (EW) indexes of NYSE stocks.

d $\beta$ = 1 is assumed for each security in forecasts.

e Forecasts are based on industry indexes alone.

f Both NYSE and industry indices are used as independent variables.

g Industry index contains the instrumental returns of the security being forecasted.

## TABLE X

### COMPARISON OF INDUSTRY AND MARKET MODELS MEAN SQUARE PREDICTION ERRORS FOR SELECTED YEARS

| | | Market | | | Market & Ind.[e] | | |
|---|---|---|---|---|---|---|---|
| Year | Naive[a] | VW | EW[b] | $\beta = 1$[c] | II[d] | II & VW | II & EW |
| 1936 | 180.15 | 148.87 | 133.91 | 142.39 | 134.41 | 140.35 | 140.91 |
| 1941 | 147.70 | 113.88 | 110.66 | 112.06 | 107.89 | 115.40 | 113.17 |
| 1946 | 111.45 | 60.47 | 59.88 | 51.53 | 54.28 | 53.42 | 52.06 |
| 1951 | 50.18 | 35.53 | 33.37 | 32.22 | 32.49 | 30.29 | 33.77 |
| 1956 | 42.35 | 33.35 | 32.30 | 30.82 | 31.56 | 32.85 | 32.23 |
| 1961 | 60.35 | 50.00 | 48.81 | 48.25 | 49.07 | 50.76 | 50.37 |
| 1966 | 66.92 | 50.66 | 48.43 | 48.19 | 48.65 | 50.22 | 49.81 |
| 1971 | 80.78 | 56.46 | 54.28 | 52.62 | 52.36 | 52.31 | 52.34 |
| 1976 | 87.11 | 58.52 | 54.55 | 53.29 | 52.90 | 51.59 | 53.36 |
| 1981 | 95.24 | 78.12 | 75.48 | 68.14 | 68.30 | 69.73 | 70.37 |
| GMSFE | 92.22 | 68.59 | 65.17 | 63.95 | 63.19 | 64.99 | 64.92 |
| SQGMSFE | 9.60 | 8.28 | 8.07 | 8.00 | 7.95 | 8.06 | 8.06 |

a Naive forecast is estimated by the mean returns for the five years prior to forecast period.

b The market indexes are the value-weighted (VW) and equally-weighted (EW) indexes of NYSE stocks.

c $\beta = 1$ is assumed for each security in conjunction with the industry index.

d The industry index consists of equally-weighted returns of 14 stocks in each industry plus the instrumental returns of the security being forecasted. The instrumental returns are computed by regressing the return of a security on the returns of the other 14 securities in the same industry.

e A value (VW) or equally weighted (EW) NYSE index and the industry index are used together to forecast returns.

**TABLE XI**

COMPARISON OF INDUSTRY AND MARKET MODELS
MEAN SQUARE PREDICTION ERRORS FOR INDUSTRIES
WITH COMPLETE DATA FROM 1936-1980

| Industry Code[a] | Naive[b] | VW | EW[c] | $\beta = 1$[d] | II[e] | II & VW[f] | II & EW[g] |
|---|---|---|---|---|---|---|---|
| 20 | 70.04 | 61.34 | 56.41 | 62.76 | 57.95 | 66.24 | 62.66 |
| 28 | 63.99 | 44.78 | 44.93 | 45.73 | 46.89 | 48.19 | 47.99 |
| 29 | 80.79 | 62.96 | 64.25 | 52.67 | 51.98 | 50.37 | 51.09 |
| 33 | 90.99 | 60.53 | 58.53 | 54.06 | 52.54 | 54.48 | 54.97 |
| 35 | 80.67 | 54.77 | 53.45 | 54.63 | 54.58 | 55.67 | 56.01 |
| 37 | 88.95 | 66.29 | 63.61 | 63.51 | 64.73 | 67.76 | 68.32 |
| Average | 80.57 | 58.45 | 56.86 | 55.56 | 54.78 | 57.12 | 56.84 |

a Industry code designates the two-digit SIC code.

b Naive forecast is estimated by the mean returns for the five years prior to forecast period.

c Market index is New York Stock Exchange (NYSE). VW and EW indicate value-weighted and equally-weighted NYSE indices respectively.

d $\beta = 1$ is assumed for each security in conjunction with the industry index.

e The industry index consists of equally-weighted returns of 14 stocks in each industry plus the instrumental returns of the security being forecasted. The instrumental variable is computed by regressing the returns of a security on the returns of the other 14 securities in the same industry group.

f,g A value (VW) or equally weighted (EW) index of NYSE stocks and the industry index are used together to forecast returns.

## TABLE XII

### SIGNIFICANCE TESTS FOR THE DIFFERENCES IN MSPE'S BETWEEN INDUSTRY AND MARKET MODELS: t-RATIOS[a]

| (1) | VW (2) | EW (3) | β = 1 (4) | II (5) | II & VW (6) | II & EW (7) |
|---|---|---|---|---|---|---|
| Naive | 16.42 | 16.42 | 15.84 | 15.35 | 14.04 | 14.13 |
| VW | | 4.78 | 6.10 | 6.04 | 3.98 | 3.77 |
| EW | | | 2.26 | 2.85 | -0.96 | 0.90 |
| B = 1 | | | | 0.49 | -2.44 | -2.14 |
| IND | | | | | -3.30 | -2.78 |
| II & VW | | | | | | 0.01 |

a t-ratios are for the differences among the GMSPE's in Table 10.

b The differences are measured between the models shown in each row of Column (1) and other models in Columns (2) through (7). t-ratio is computed by $(\bar{d}/\sigma_d)\sqrt{T}$, where $d = (1/T)(SPE_t^{(i)} - SPE_t^{(j)})$ and $\sigma_d = (1/T)\sum_{t=1}^{T}(d_t - \bar{d})^2$. $SPE_t^{(i)}$ is the squared forecast error for the ith model in each column.