On the Robustness of Goodness-of-Fit Criteria for Factor Identification: Simulation and Some Korean Evidence

Sang Bin Lee
Korea Advanced Institute of Science and Technology

Tae Yol Yoo
Korea Advanced Institute of Science and Technology

Abstract

Recent empirical studies on the arbitrage pricing theory (APT) have focused on the factor identification problem for securities returns. Using simulated returns generated under violations of the assumptions of the factor model, this paper compares goodness-of-fit criteria based on maximum likelihood analysis by changing the number of factors, the number of securities, and the number of observations. For a variety of simulation designs, Schwartz's Bayesian criterion is robust to all violations of the assumptions in the factor model. The cross validation performs well for a small number of securities, but its robustness is limited to only a large sample as the number of securities increases. Schwartz's Bayesian criterion identifies three factors using actual returns in the Korean stock market.

Introduction

A recent issue in empirical tests of the arbitrage pricing theory developed by Ross (1976) is factor identification. Roll and Ross (1980), Dhrymes, Friend, and Gultekin (1984), Cho and Taylor (1987), and Gultekin and Gultekin (1987) use the standard likelihood ratio (LR) statistic to estimate the number of factors under the exact factor structure, while Trzcinka (1986) and Brown (1989) use the eigenvalue test under the approximate factor structure. Brown (1989) points out that the eigenvalue test of factor identification has serious limitations; in a simple economy where there are only k equally important factors, eigenvalue analysis may lead one to infer falsely that only one factor is responsible for securities returns. This paper focuses on factor identification by the LR statistic in a maximum likelihood (ML) estimation.

Several financial researchers have cautioned against use of the formal LR statistic in financial applications because several violations of the assumptions of the standard factor model typically occur with actual return data. Brown and Weinstein (1985), for example, indicate that the residuals from daily stock returns are negatively autocorrelated and highly nonnormal. Perry (1982) shows that if daily stock returns are used, the serial independence assumption may not

0747-5535/92/1500-0018 $01.00

be valid. Thus, serial correlation and nonnormality frequently occur in applications of factor analysis that use actual stock returns.

A second violation may occur in empirical tests of APT. The residual portion of the covariance matrix is not a diagonal matrix and thus can be cross correlated; the residual factor may represent industry or firm size factors. Connor and Korajczyk (1988) use cross-correlated returns for simulation of an asset return series that conforms to an approximate factor model.

Another unrealistic assumption in exploratory factor analysis is that common factors are uncorrelated. In many financial applications, it is more natural to assume that economic factors are correlated. Conway and Reinganum (1988) simulate the correlated two factor structure with prior information in order to assess the robustness of factor identification.

To identify the correct number of factors under the above violations of the assumptions of the standard factor model, several goodness-of-fit criteria can be considered as alternatives to the strict test. This paper examines cross validation (CV), Akaike information criterion (AIC), and Schwartz's Bayesian criterion (SBC).

CV is a general statistical model identification method that determines whether an estimated model reflects stable features of the underlying process. It can be used to identify the stable number of factors by fitting successive models with additional factors and noting when the prediction errors begin to stabilize or increase. Large samples are needed for the CV index because the data are divided into two halves, i.e., a fitted sample and a validation sample.

Conway and Reinganum (1988) use the CV technique to identify the stable number of factors for securities returns. They find that there is one dominant factor and one minor factor that explain securities returns, regardless of the group of securities analyzed. These results contrast with other empirical evidence produced by the standard LR test (e.g., Roll and Ross (1980) find between three and five factors; Dhrymes, Friend, and Gultekin (1984) find more than five factors; and Cho and Taylor (1987) find between six and seven factors, etc.).

Cudeck and Brown (1983), however, argue that AIC and SBC may be used in place of CV for small samples. AIC and SBC include a penalty function that depends on the number of parameters fitted. SBC is an asymptotic version of a Bayes procedure that assigns positive probability to lower dimensional subspaces of the parameter space, while AIC is used to select a model that minimizes the loss function through information theory. As more parameters are added to a model, the decrease in the LR statistic must be large enough to warrant the increase in the number of fitted parameters. These two measures seek to insure that the decrease in the LR statistic caused by fitting additional parameters is large enough to justify the larger model. Elton and Gruber (1988) use SBC to estimate the factor structure in the Japanese stock market. Jobson compares the

CV index with AIC and SBC in his comment on Conway and Reinganum (1988).

This paper evaluates the robustness of goodness-of-fit criteria based on LR statistics under four violations of the assumptions of the standard factor model: serial and cross correlation, nonnormality, and correlated factor structure. Using simulated returns generated under violations of the assumptions, the performances of goodness-of-fit criteria are compared. To assess the robustness of goodness-of-fit criteria under all experimental environments, a simulation also is run with changes in the number of factors, the number of securities, and the number of observations. For a variety of simulation designs, SBC is robust to all violations of the assumptions in the factor model. CV performs well for a small number of securities. As the number of securities increases, however, its robustness is limited to only a large sample. Factor identification also is performed based on the goodness-of-fit criteria using actual daily returns in the Korean stock market; SBC identifies three factors.

Design of Simulation Experiments

Violations of the Assumptions of the Factor Model

It is assumed that a general k factor linear model underlies the return-generating process. Let R_{it} represent the observed stock return and μ_i the corresponding expected return on the ith security (i = 1, 2, ... , n) at time period t. The standard factor model is written as:

$$R_{it} = \mu_i + b_{i1}f_{1t} + b_{i2}f_{2t} + \ldots + B_{ik}f_{kt} + \varepsilon_t$$

where:

- f_{jt} = jth risk factor that impacts asset returns, where j = 1, 2, ... , k are k different risk factors. The jth risk factor has a mathematical expectation of zero, $E(f_{jt}) = 0$;
- b_{ij} = A factor loading (or sensitivity) indicator that measures how responsive returns from asset i are to index j for j = 1, 2, ... , k;
- ε_{it} = A random error (or residual) term for asset i in period t that measures unexplained residual return; it has an expected value of zero and a variance $\sigma_i^2(\varepsilon_{it})$.

The market model can be considered as a special case of the model where there is only one factor—the return on a market portfolio. When the factors f_{jt}

and residuals ε_{it} are multivariate normal and uncorrelated, factor analysis can be used to obtain ML estimates of the factor. But serial correlation and nonnormality of actual stock returns raise doubts about the usefulness of the ML factor analysis procedure which assumes that the residual terms are normal and independent across time. The following four violations of the assumptions of the standard factor model are considered in assessing the robustness of goodness-of-fit criteria for factor identification.

- Serial correlation;
- Cross correlation;
- Correlated factor structure;
- Nonnormality.

The first violation of the models assumptions tested is serial correlation. To examine the performance of goodness-of-fit criteria, it is assumed that the residual terms follow a first order autoregressive (AR(1)) process, i.e., $\varepsilon_{it} = \rho\ \varepsilon_{i,t-1} + a_{it}$, where a_{it} has a normal distribution with a mean 0 and variance $\sigma_i^2(a_{it})$. By changing the correlation level, the robustness of the goodness-of-fit criteria is assessed according to the magnitude of serial correlation.

The second violation that may occur in the standard factor model is cross correlation of stock returns. This paper considers cross correlation of the residual terms estimated from the factor model, i.e., correlation between the residual terms of ith stock and jth stock. The simulated residuals are generated from a multivariate normal distribution [0,V], where V represents the general covariance matrix of the residual terms having the correlation level, ρ.

Another unrealistic assumption in the exploratory factor is the assumption that common factors are correlated. This assumption is unrealistic in many financial applications, in which it may be more natural to assume that the economic factors are correlated. Therefore, this research considers the case in which the factors are not independent, but correlated. It is assumed that the factors have multivariate normal distribution $[0,\psi]$, where ψ represents the general covariance matrix with factor correlation ρ.[1] Similarly, for the case of serial correlation, the

[1] To generate the general covariance matrix V of the residual terms having the correlation level ρ, the elements of covariance matrix v_{ij} are determined by

$$v_{ij} = \rho\, v_i\, v_j,$$

where v_i is given by $\sigma_i\,(\varepsilon_{it})$.

Similarly, for the general covariance matrix ψ of the factor,

$$\psi_{ij} = \rho\, \psi_i\, \psi_j$$

where ψ_i is the standard deviation of factor.

level of correlation is increased to assess the robustness of goodness-of-fit under cross correlation and under the correlated factor structure.

Finally, to examine the robustness of goodness-of-fit criteria under nonnormality, the method of marginal moments used by Affleck-Graves and MacDonald (1989) is employed. Fleishman (1978) presents a method for generating univariate nonnormal errors with the specified first four moments (mean, variance, skewness, and kurtosis for each marginal distribution), and Vale and Maurelli (1983) extend the Fleishman procedure to generate multivariate nonnormal variables with a specified covariance structure and specified first four marginal moments.[2] Affleck-Graves and McDonald (1989) apply this method to determine the effects of nonnormalities on the multivariate test. This paper uses the method of marginal moments to generate the nonnormal residual terms for the simulation.

Description of Some Criteria

In exploratory factor analysis, identification of the correct number of factors is a difficult problem. Various statistical tests of the residual covariance matrix after a k factor model has been fit to the data have been suggested. A conventional goodness-of-fit measure for identifying the number of factors is the standard LR statistic proposed by Lawley and Maxwell (1963). This statistic has the form:

$$LR = T^* (\log |BB' + D| - \log |S|) + T^*(\mathrm{tr}[BB' + D]^{-1}S - n)$$

where:

---

[2] For a nonnormal random variable Y, they utilize knowledge of the first four moments of the power transformation,

$$Y = a + bX + cX^2 + dX^3$$

where:

X	=	A standardized normal deviate; and
a, b, c, and d	=	Constants that can be selected so that Y has the desired first four moments.

By evaluating the skewness and kurtosis of Y, a system of four nonlinear equations with four unknowns can be used to determine a, b, c, and d as seen in equations (2), (3), (4), and (5) of Vale and Maurelli (1983). If $\gamma_{y1,y2}$ is the correlation between the nonnormal variable Y_1 and Y_2 and $\rho_{x1,x2}$ is the corresponding correlation of the normal variates, then they show that solving the polynomial

$$\gamma_{y1,y2} = \rho_{x1,x2}(b_1b_2 + 3b_1d_2 + 3d_1b_2 + 9d_1d_2) + \rho^2_{x1,x2}(2c_1c_2) + \rho^3_{x1,x2}(6d_1d_2)$$

for $\rho_{x1,x2}$ provides the correlation between X1 and X2 required to generate the desired posttransformation correlation $\gamma_{y1,y2}$.

$T^* = T - \frac{2n + 4k + 11}{6}$ = Bartlett's corrected coefficient in moderate sample size;

B and D = ML estimates of factor loading and residual variance; and

S = Sample covariance matrix.

The LR statistic measures the overall error in factor estimates of the sample covariances by comparing the generalized variances, $|BB' + D|$ and $|S|$. When the normality assumptions for factors and residuals apply, the LR statistic has an asymptotic central chi-squared distribution with $[(n - k)^2 - (n - k)]$ degrees of freedom. Several authors have suggested using the LR statistic as a guideline rather than as a definitive statement about the number of factors, however, because violations of the assumptions typically occur with real data.

As an alternative to the standard LR statistic, three goodness-of-fit measures for identifying the number of factors are considered in this paper. The CV index is the LR value obtained by replacing S of the fitted sample with S* of the validation sample. The relation between the LR statistic with AIC and SBC is:

$$AIC = LR + 2q$$
$$SBC = LR + q \log T^*$$

where:

$$q = \frac{n(k + 1) - k(k - 1)}{2} = \text{The number of parameters fitted.}$$

As a penalty value, AIC gives the value of multiplying the number of parameters fitted by two. SBC gives the value of multiplying the number of parameters fitted by log T*. Because the number of periods, T, is large relative to the number of securities, n, and the number of factors, k, the value of log T* is greater than two.[3] This means that as additional factors are added, the penalty value of SBC is larger than that of AIC.

[3]The marginal increases of AIC and SBC over the introduction of additional factors are:

$$\frac{\partial AIC}{\partial k} = \frac{\partial LR}{\partial k} + 2 \times \frac{n - k + 1}{2}$$

$$\frac{\partial SBC}{\partial k} = \frac{\partial LR}{\partial k} + \log \frac{T - (2n + 4k + 11)}{6} \times \frac{(n - k + 1)}{2} +$$

$$\frac{-4}{6T - (2n + fk + 11)} \times \frac{n(k + 1) - k(k - 1)}{2}$$

Design of the Simulation Experiment

The parameters for the simulation experiment[4] are estimated with 1,171 daily returns from the KIS-KAIST Stock Price System (January 1985 to December 1988). First, factor scores are estimated from the ML estimation using factor analysis. Then factor loadings and variances of residuals are estimated by regressing securities returns on the time series of these factor realizations.[5] The parameters are fixed in the simulations for violations of the assumptions. For the serial correlation, the simulated residuals from the AR(1) model are generated varying the correlation level. Finally, for cross correlation and the correlated factor structure, residual terms and factors are generated from the general multivariate normal distributions using the RNMVN function in the IMSL language.

Under each violation of the assumptions of the standard factor model, 100 simulated samples are generated with simulated returns for each security. For the CV index, the even-numbered returns in each sample are used to estimate the factor loading and residual variance, and the odd-numbered returns are used to validate the estimated model. AIC and SBC use simulated returns to calculate the value of criteria.

To examine the robustness of goodness-of-fit criteria under the three correlation structures, the simulation is designed in three ways:

- Robustness of goodness-of-fit criteria with changes in the number of factors;
- Robustness of goodness-of-fit criteria with the nested sequence of securities; and
- Robustness of goodness-of-fit criteria with an increase in the number of observations.

The first concern is whether the number of factors affects the robustness of goodness-of-fit criteria. For the simulation, one, three, and five factor models

---

Because the third term in the latter is approximately zero when T is large relative to n and k, the marginal penalty values of AIC and SBC depend on both 2 and log (T - (2n + 4k +11)/6).

[4]This system has been developed jointly by the Korean Advanced Institute of Science and Technology and the Korea Investors' Services. It includes all return data from 1980 through 1988 in the Korean stock market.

[5]In general there are two approaches in the empirical testing of APT. Most of the initial APT empirical researchers have used factor analysis to estimate the risk factors and the factor loadings endogenously (Roll and Ross, 1976). Risk factors that may affect security prices can be developed from exogenous economic theory, however, as well as from endogenous factor analytic techniques. As an alternative to using factor analysis to estimate an APT model, some financial analysts and economists (Chen, Roll, and Ross, 1986) prefer to estimate the APT model using a different empirical testing procedure—a two step procedure somewhat like the procedure used to estimate the CAPM empirically. For this simulation experiment, the risk factor is estimated endogenously from factor analysis. Cross sectional regression then is used to estimate the factor loadings or the coefficients of the regression model.

with 30 fixed securities and 1,200 observations are used to assess the performances of the goodness-of-fit criteria. The parameters of factor scores f_t are estimated from the ML estimation for each model.[6] Then a cross-sectional regression is applied to estimate the factor-loading coefficients b_{ij} and residual variance $\sigma_i^2(\varepsilon_{it})$.

Second, Dhrymes, Friend, and Gultekin (1984) show that as the number of securities in a group increases, the number of factors determined by the standard LR statistic increases. The behavior of the number of factors identified by goodness-of-fit criteria when the number of securities increases is the focus of interest in this research. This paper assesses the performance of goodness-of-fit criteria using the nested sequence approach for the number of securities where a smaller size sample of simulated returns is a subset of the larger size sample.

Shukla and Trzcinka (1990) and Trzcinka (1986) use both the nested and the random sequence of sample covariance matrices to compare factor analysis and principal component analysis.[7] They show that the results are similar for both sequences. Hence, this paper employs the nested sequence that grows by adding rows and columns to the smallest sample matrix. Simulated returns with 30, 60, and 90 securities by the nested sequence are generated as fixed with three factors and 1,200 observations under each violation of the standard factor model.

Third, AIC and SBC originally were introduced for the small sample that may be used in place of CV which is for a large sample. According to sample size, therefore, AIC and SBC should be compared to CV to see if the performances are still comparable. In order to examine the effects of goodness-of-fit criteria with the change of sample size, this paper changes the number of observations to 1,200, 2,000, and 3,000 for the factor model having three factors and 60 securities.

Finally, to examine the behavior of goodness-of-fit criteria under nonnormality, simulated returns having the multivariate nonnormal distributions that are characterized by the first four moments are generated by the three factor model with 30 securities and 1,200 observations. To see whether the magnitude of nonnormality affects the robustness of goodness-of-fit criteria, this paper runs the simulation varying the value of skewness and kurtosis. Two separate effects are considered for the robustness of goodness-of-fit criteria under nonnormality,

[6] For the one factor model, f_1 has a normal distribution with zero standard deviation of .919838. For the three factor model, f_1, f_2, and f_3 are normally distributed with zero mean and standard deviations of .9225, .8045, and .5305. Similarly, the five factor model has standard deviations of .9252, .8114, .5469, .4898, and .3977.

[7] Principal component analysis has several advantages over factor analysis: uniqueness up to scalar transformation, likelihood to reject the null hypothesis, and ease in computation. Shukla and Trzcinka (1990) argue that principal component analysis may be preferable to factor analysis when the size of the covariance matrix (sample size = 30, 33, 36, ... , 596) grows larger, thus allowing isolation of the theoretical pricing error.

i.e., the net effect of nonnormality and the joint effect of nonnormality and the intermediate correlation—cross correlation between the residuals and ith stock and jth stock.

Simulation Results

Results Under the Correlation Structures

To assess the robustness of goodness-of-fit criteria under each correlation structure, the simulation is run varying the number of factors, the number of securities, and the number of observations. Table 1 shows the behavior of goodness-of-fit criteria for the standard three factor model having no correlation with 30 securities and 1,200 observations. The values of the LR statistic and goodness-of-fit criteria ranging between one and ten factors, averaged over 100 simulated samples, are presented in Table 1. The values of CV, AIC, and SBC decline sharply as the number of factors moves from one to two. For example, the value of SBC drops from 6,406 to 2,340. The values continue to decrease monotonically until a third factor is added to the model and gradually increase as additional factors are added. Hence, all criteria for the three factor model with no correlation structure identify three factors.

The average LR statistic continues to decline monotonically as additional factors are added to model. The ρ values of Table 1 show that the LR statistic does not identify the correct number of factors, even though the standard model with no correlations is considered. This result confirms why several authors have used the LR statistic as a guideline rather than as a definitive statement identifying the number of factors. Hereafter, the focus of this paper is the problems of factor identification by CV, AIC, and SBC, excluding the LR statistic.

Figure 1 graphs the pattern of goodness-of-fit criteria in the case of the standard three factor model. Except for the LR value, the values of the other criteria are minimal at the third factor. The values of SBC show a more clear pattern than do the values of AIC. In the design of the simulation, the value of log T^* is approximately 7. The marginal increases of AIC and SBC over the introduction of additional factors are 2q and 7q. Because SBC has an additional penalty 5q, it is natural that it has a clearer pattern than does AIC.

Next the behavior of goodness-of-fit criteria are examined under each correlation structure by increasing the values of correlations by 0.05 over the 0.05 through 0.50 range for serial and the cross correlations and by 0.1 from 0.1 through 1.0 for the correlated factor structure. Table 2 shows the number of factors identified under each correlation structure for the three factor model with 30 securities and 1,200 observations. In the serial correlation, CV and SBC identify the correct number of factors (through the highest level of correlation considered in this paper). But AIC seems to overestimate the number of factors when the level of correlation exceeds a value of 0.20. The result is similar to the cross correlation. In the cross correlation, the maximum levels of correlation of CV, AIC, and SBC identifying the correct number of factors are 0.50, 0.15, and 0.50.

In the correlated factor structure, all factors must be integrated into one factor if the level of correlation approaches 1. Table 2 shows that the number of factors identified at the correlation level $\rho = 1$ is one for all goodness-of-fit criteria. Table 2 shows that all criteria identify the correct number of factors even at the correlation level $\rho = 0.7$ to 0.9 and thus perform well for the correlated factor structure. The results for the three factor model with 30 securities and 1,200 observations show that AIC tends to overestimate the number of factors with serial and cross correlation, while CV and SBC are robust to three violations of the correlation structure.

Table 3.1 presents the maximum level of correlation identified in each correlation structure for one, three, and five factor models with 30 securities and 1,200 observations. The maximum level of correlation identifying the correct number of factors for the three factor model of Table 2 is summarized in Panel B of Table 3.1. Panels A and B of Table 3.1 show that the results for the one and five factor models are similar to those of the three factor model. CV and SBC are robust for all correlation structures, but AIC does not identify the correct number of factors. Table 3.1 shows that the performance of the goodness-of-fit criteria are not dependent on the number of factors.

Table 3.2 assesses the performance of goodness-of-fit criteria in each correlation structure using the nested sequence approach for the number of securities. The number of securities is varied (30, 60, or 90). The simulated sample of 30 securities is a subset of the 60 securities sample, which is also a subset of the 90 securities sample. As the number of securities increases (moving from Panel A to Panel B to Panel C), the maximum level of correlation identified with each correlation structure decreases sharply. For example, the correlation level of CV in the serial correlation drops from 0.50 to 0.15 when the number of securities increases from 30 to 60. As the number of securities increases to 90, CV does not identify the correct number of factors even at $\rho = 0.05$ (the lowest level considered in the current simulation). Other criteria perform similarly to CV, except for SBC in the serial correlation. SBC performs well for serial correlation, irrespective of the number of securities.

In the above experiment, the number of observations becomes relatively small as the number of securities increases. This means that the decrease of performance of goodness-of-fit criteria with the number of securities may be due to the small sample size. To resolve this problem, the simulation is run again. The number of observations is increased, while the number of securities is fixed at 60. Table 3.3 presents the summary values for the correlation level with increases in the number of observations from 1,200 to 2,000 and to 3,000. As the number of observations increases, performances of all the criteria except the CV index seem to remain unchanged. The CV index in the serial correlation improves rapidly—the maximum level of correlation increases from 0.05 in Panel A to 0.45 in Panel B and to 0.50 in Panel C. Performance seems to become a little better as the number of observations increases for the CV index in the cross correlation. As the sample size increases, therefore, CV performs well and the other criteria remain unchanged.

The results under the correlation structures together yield several findings:

- For the factor model with 30 securities and 1,200 observations, CV and SBC are robust to both the correlation structure and number of factors. AIC overestimates the number of factors in the serial and cross correlation and is sensitive to change in the number of factors;
- When the number of securities with fixed small samples is increased, the robustness of CV drops substantially. On the other hand, SBC still performs well, regardless of the number of securities;
- As the number of observations increases, the performance of CV in the serial correlation improves sharply.

Results Under Nonnormality

To assess the robustness of goodness-of-fit criteria according to the magnitude of nonnormality, this paper runs the simulation experiment varying the value of skewness (γ_1 = 0, 1, 2) and kurtosis (γ_2 = 0, 2, 4, 6) simultaneously. Panel A in Table 4 shows the number of factors identified by the goodness-of-fit criteria under a structure of nonnormality and no correlation. All goodness-of-fit criteria correctly identify three factors at all levels of skewness and kurtosis in the simulation. Hence, all criteria are robust to the violation of pure nonnormality with no correlation. The results are comparable with those of Affleck-Graves and Mcdonald (1989) who argue that the multivariate test is reasonably robust with respect to typical levels of nonnormality.

The other concern related to the nonnormality problem is the behavior of goodness-of-fit criteria when nonnormality is combined with the correlation structure. Another simulation is run to examine the question of nonnormality and the intermediate correlation. The simulation results with correlation levels of 0.1, 0.2, and 0.3 are presented in Panels B, C, and D of Table 4. At the correlation level of 0.1, all criteria (except the $\gamma_1 = 2$, $\gamma_2 = 6$ of SBC) identify the correct number of factors. As the correlation level increases, Panels C and D show that AIC overestimates the number of factors at all levels of skewness and kurtosis, while CV and SBC identify the correct number of factors. These findings are consistent with the behavior of goodness-of-fit criteria for the factor model having no nonnormality under the correlation structure (Table 2). Hence, the robustness of goodness-of-fit criteria is not affected by departures from the normality of stock returns, but is affected by nonnormality combined with intermediate correlation according to the magnitude of correlation level.

Empirical Evidence in the Korean Stock Market

This section applies goodness-of-fit criteria to daily returns in the Korean stock market. The data used are obtained from the second version of the KIS-KAIST Stock Price System. This study considers 210 firms with 1,171 daily return data from January 1985 through December 1988.[8] Securities are divided into seven groups of 30 securities. A group is selected first by ranking all stocks by industry code in the Korean stock market system and assigning every seventh stock to a particular group.

For each of the seven groups, the values of the LR statistic and goodness-of-fit criteria with one to ten factors are obtained using 1,171 trading days. The value of CV is obtained using 586 odd trading days as the fitted sample and validating with 585 even trading days. Table 5 displays the values of these criteria for models of zero through ten factors averaged over the seven groups. CV decreases sharply until the fifth factor is added and stabilizes as additional factors are added. CV identifies five factors. Also, AIC is minimal at the sixth factor, while SBC is minimal at the third factor. In actual data, SBC identifies a smaller number of factors than does CV or AIC. It is interesting to compare with the Japanese results of Elton and Gruber (1988) who report that there are four factors present in the return-generating process using SBC for the Japanese stock market.

Conclusion

Several financial researchers caution against use of the formal LR statistic in financial applications because of violations of the assumptions in actual data. This paper considers four violations of the assumptions: serial correlation, cross correlation, correlated factor structure, and nonnormality. For the simulated returns generated under violations of the assumptions of the factor model, goodness-of-fit criteria based on the maximum likelihood analysis are compared by changing the number of factors, the number of securities, and the number of observations. The findings are:

[8] The 210 firms were selected from a total 323 firms by order of number of less than zero returns.

- Under all correlation structures, the CV and SBC are indifferent to change in the number of factors, while AIC is sensitive to the number of factors;
- When the number of securities in a group with a fixed small sample is increased, the power identifying the correct number of factors in CV (specifically, for the serial correlation) drops substantially. On the other hand, SBC performs well irrespective of the number of securities;
- As the sample size (number of observations) increases, the performance of CV in the serial correlation improves sharply;
- All criteria are robust to the pure nonnormality with no correlation. But for nonnormality that has an intermediate correlation, CV and SBC perform well.

To summarize, SBC seems to be robust to all violations of assumptions of the factor model considered in this paper. CV performs well when a small number of securities is used. As the number of securities increases, however, its robustness is limited to only a large sample.

Finally, factor identification is performed based on goodness-of-fit criteria using actual daily returns in the Korean stock market. In this instance, SBC was found to identify three factors.

References

1. Affleck-Graves, John, and Bill Mcdonald, "Nonnormality and Tests of Asset Pricing Theories," *Journal of Finance*, 44 (September 1989), pp. 889-908.

2. Akaike, H., "A New Look at the Statistical Model Identification," *IEEE Transactions on Automatic Control*, 19 (December 1974), pp. 716-723.

3. Brown, S.J., "The Number of Factors in Security Returns," *Journal of Finance*, 44 (December 1989), pp. 1247-1262.

4. Brown, S.J., and M.I. Weinstein, "Derived Factors in Event Studies," *Journal of Financial Economics*, 14 (September 1985), pp. 491-495.

5. Chen, Nai-fu, R. Roll, and S.A. Ross, "Economic Forces and the Stock Market: Testing the APT and Alternative Pricing Theories," *Journal of Business*, 59 (1986), pp. 491-495.

6. Cho, D.C., and W.M. Taylor, "The Seasonal Stability of the Factor Structure of Stock Returns," *Journal of Finance*, 5 (December 1987), pp. 1195-1211.

7. Connor, G., and R.A. Korajczyk, "Risk and Return in an Equilibrium APT," *Journal of Financial Economics*, 21 (1988), pp. 255-289.

8. Conway, D.A. and M.R. Reinganum, "Stable Factors in Security Returns: Identification Using Cross-Validation," *Journal of Business & Economic Statistics*, 6 (April 1988), pp. 1-15.

9. Cudeck, R., and M.W. Browne, "Cross-Validation of Covariance Structures," *Multivariate Behavioral Research*, 18 (April 1983), pp. 147-167.

10. Dhrymes, P., I. Friends, and B. Gultekin, "A Critical Reexamination of the Empirical Evidence on the Arbitrage Pricing Theory," *Journal of Finance*, 39 (June 1984), pp. 323-346.

11. Elton, E.J., and M.J. Gruber, "A Multi-Index Risk Model of the Japanese Stock Market," *Japan and the World Economy*, 1 (1988), pp. 21-44.

12. Fleishman, Allen I., "A Method of Simulating Nonnormal Distributions," *Psychometrica*, 43 (December 1978), pp. 521-532.

13. Gultekin, M., and N. Gultekin, "Stock Return Anomalies and the Tests of the APT," *Journal of Finance*, 42 (December 1987), pp. 1213-1224.

14. Jobson, J.D., "Stable Factors in Security Returns: Identification Using Cross-Validation: Comment," *Journal of Business & Economic Statistics*, 6 (January 1988), pp. 16-20.

15. Lawley, D.N., and A.E. Maxwell, *Factor Analysis as a Statistical Method* (London: Butterworth, 1963).

16. Perry, P., "The Time-Variance Relationship of Security Returns: Implications for the Return-Generating Stochastic Process," *Journal of Finance*, 37 (June 1982), pp. 858-869.

17. Roll, R., and S.A. Ross, "An Empirical Investigation of the Arbitrage Pricing Theory," *Journal of Finance*, 36 (December 1980), pp. 1073-1103.

18. Ross, S.A., "The Arbitrage Theory of Capital Asset Pricing," *Journal of Economic Theory*, 13 (1976), pp. 341-360.

19. Schwartz, G., "Estimating the Dimension of a Model," *The Annals of Statistics*, 6 (1978), pp. 341-360.

20. Shukla, R., and C. Trzcinka, "Sequential Tests of the Arbitrage Pricing Theory: A Comparison of Principal Components and Maximum Likelihood Factors," *Journal of Finance*, 45 (December 1990), pp. 1541-1564.

21. Trzcinka, C., "On the Number of Factors in the Arbitrage Pricing Model," *Journal of Finance*, 2 (June 1986), pp. 347-368.

22. Vale, C. David, and Vincent A. Maurelli, "Simulating Multivariate Nonnormal Distributions," *Psychometrica*, 48 (September 1983), pp. 465-471.

Table 1
Summary Values of Goodness-of-Fit Criteria for 100 Simulated Samples for the Standard Model With No Correlation (Three Factor Model With 30 Securities and 1,200 Observations)

Number of Factors	LR	CV	AIC	SBC
1	5,981 (.00)	4,690	6,101	6,406
2	1,710 (.00)	1,671	1,888	2,340
3	413 (.00)	580*	647*	1,241*
4	363 (.05)	666	651	1,382
5	322 (.13)	759	662	1,525
6	289 (.20)	865	679	1,669
7	261 (.24)	972	699	1,811
8	238 (.23)	1,075	722	1,950
9	220 (.17)	1,171	748	2,087
10	204 (.11)	1,269	774	2,221

ρ-values greater than chi-squared values with each degree of freedom are shown in parentheses
* Number of factors identified in the standard model is three

Table 2
Summary Values for the Number of Factors Identified on Goodness-of-Fit Criteria Under Each Correlation Structure (Three Factor Model With 30 Securities and 1,200 Observations

Level of Correlation	CV	AIC	SBC
		Serial Correlation	
$\rho = 0.05$	3	3	3
$\rho = 0.10$	3	3	3
$\rho = 0.15$	3	3	3
$\rho = 0.20$	3	3	3
$\rho = 0.25$	3	4	3
$\rho = 0.30$	3	4	3
$\rho = 0.35$	3	4	3
$\rho = 0.40$	3	5	3
$\rho = 0.45$	3	6	3
$\rho = 0.50$	3	7	3
		Cross Correlation	
$\rho = 0.05$	3	3	3
$\rho = 0.10$	3	3	3
$\rho = 0.15$	3	3	3
$\rho = 0.20$	3	4	3
$\rho = 0.25$	3	4	3
$\rho = 0.30$	3	4	3
$\rho = 0.35$	3	4	3
$\rho = 0.40$	3	4	3
$\rho = 0.45$	3	4	3
$\rho = 0.50$	3	4	3
		Correlated Factor	
$\rho = 0.10$	3	3	3
$\rho = 0.20$	3	3	3
$\rho = 0.30$	3	3	3
$\rho = 0.40$	3	3	3
$\rho = 0.50$	3	3	3
$\rho = 0.60$	3	3	3
$\rho = 0.70$	3	3	3
$\rho = 0.80$	3	3	2
$\rho = 0.90$	2	3	2
$\rho = 1.00$	1	1	1

Table 3.1
Summary Values for Level of Correlation With Change in Number of Factors (One, Three, and Five Factor Model With 30 Securities and 1,200 Observations)

	CV	AIC	SBC
	Panel A One Factor Model*		
Serial Correlation	0.50	0.30	0.50
Cross Correlation	0.15	0.05	0.35
	Panel B Three Factor Model		
Serial Correlation	0.50	0.20	0.50
Cross Correlation	0.50	0.15	0.50
Correlated Factor	0.90	0.90	0.70
	Panel C Five Factor Model		
Serial Correlation	0.50	--**	0.50
Cross Correlation	0.35	--**	0.50
Correlated Factor	0.70	--**	0.60

* Simulation under correlated factor structure is excluded because the one factor model cannot identify more than one factor

** AIC does not identify the correct number of factors, even at the lowest correlation level considered in this paper

Table 3.2
Summary Values for Level of Correlation With the Nested Sequence of the Number of Securities (Three Factor Model With 30, 60, and 90 Securities and 1,200 Observations)

	CV	AIC	SBC
	Panel A 30 Securities		
Serial Correlation	0.50	0.20	0.50
Cross Correlation	0.50	0.15	0.50
Correlated Factor	0.80	0.90	0.70
	Panel B 60 Securities		
Serial Correlation	0.15	0.20	0.50
Cross Correlation	--**	--**	0.15
Correlated Factor	0.90	0.90	0.80
	Panel C 90 Securities		
Serial Correlation	--**	0.15	0.50
Cross Correlation	--**	--**	0.15
Correlated Factor	--**	0.90	0.80

** AIC does not identify the correct number of factors, even at the lowest correlation level considered in this paper

Table 3.3
Summary Values for Level of Correlation With the Increases of the Number of Observations (Three Factor Model With 60 Securities and 1,200, 2,000 and 3,000 Observations)

	CV	AIC	SBC
	Panel A 1,200 Observations		
Serial Correlation	0.15	0.20	0.50
Cross Correlation	--**	--**	0.15
Correlated Factor	0.90	0.90	0.80
	Panel B 2,000 Observations		
Serial Correlation	0.45	0.20	0.50
Cross Correlation	0.05	--**	0.10
Correlated Factor	0.90	0.90	0.80
	Panel C 3,000 Observations		
Serial Correlation	0.50	0.25	0.50
Cross Correlation	0.05	--**	0.10
Correlated Factor	0.90	0.90	0.80

** AIC does not identify the correct number of factors, even at the lowest correlation level considered in this paper

Table 4
Summary Values for the Number of Factors Identified on Goodness-of-Fit Criteria Under Nonnormality (Three Factor Model With 30 Securities and 1,200 Observations)

γ_1 - Skewness γ_2 - Kurtosis	CV	AIC	SBC	CV	AIC	SBC
		Panel A $\rho = 0.00$			Panel B $\rho = 0.10$	
$\gamma_1 = 0, \gamma_2 = 0$	3	3	3	3	3	3
$\gamma_1 = 0, \gamma_2 = 2$	3	3	3	3	3	3
$\gamma_1 = 0, \gamma_2 = 4$	3	3	3	3	3	3
$\gamma_1 = 0, \gamma_2 = 6$	3	3	3	3	3	3
$\gamma_1 = 1, \gamma_2 = 2$	3	3	3	3	3	3
$\gamma_1 = 1, \gamma_2 = 4$	3	3	3	3	3	3
$\gamma_1 = 1, \gamma_2 = 6$	3	3	3	3	3	3
$\gamma_1 = 2, \gamma_2 = 2$	3	3	3	3	3	3
$\gamma_1 = 2, \gamma_2 = 4$	3	3	3	3	3	3
$\gamma_1 = 2, \gamma_2 = 6$	3	3	3	3	4	3
		Panel C $\rho = 0.20$			Panel D $\rho = 0.30$	
$\gamma_1 = 0, \gamma_2 = 0$	3	4	3	3	4	3
$\gamma_1 = 0, \gamma_2 = 2$	3	4	3	3	4	3
$\gamma_1 = 0, \gamma_2 = 4$	3	4	3	3	4	3
$\gamma_1 = 0, \gamma_2 = 6$	3	4	3	3	4	3
$\gamma_1 = 1, \gamma_2 = 2$	3	4	3	3	4	3
$\gamma_1 = 1, \gamma_2 = 4$	3	4	3	3	4	3
$\gamma_1 = 1, \gamma_2 = 6$	3	4	3	3	5	3
$\gamma_1 = 2, \gamma_2 = 2$	3	4	3	3	5	3
$\gamma_1 = 2, \gamma_2 = 4$	3	4	3	3	5	3
$\gamma_1 = 2, \gamma_2 = 6$	3	5	3	3	6	3

Table 5
Summary Values for Goodness-of-Fit Criteria of Seven Randomly Selected Securities in Korean Data (January 1985 through December 1988)

Number of Factors	LR	CV	AIC	SBC
1	2,378	2,432	2,498	2,759
2	1,106	1,522	1,284	1,671
3	806	1,339	1,040	1,549
4	643	1,289	931	1,557
5	543	1,245	883	1,622
6	482	1,241	872	1,719
7	438	1,240	876	1,827
8	404	1,230	888	1,940
9	379	1,230	907	2,054
10	363	1,218	933	2,170

Figure 1
Simulation Result
Standard Three Factor Model

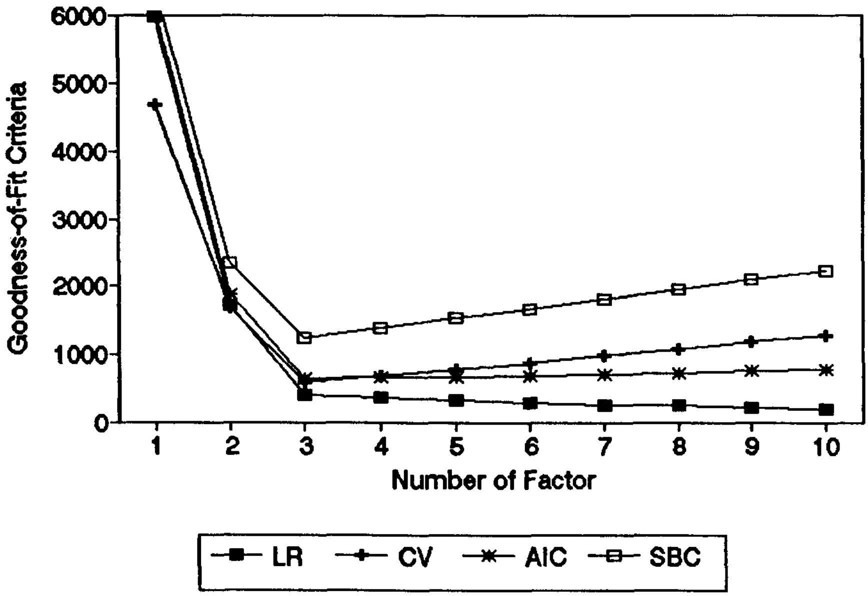

Figure 2.a
Number of Factors
Serial Correlation

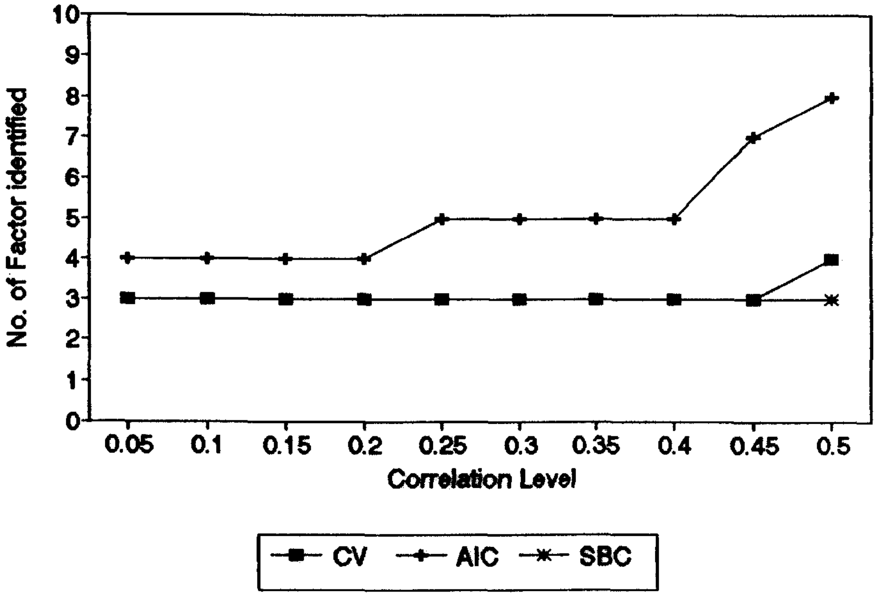

Figure 2.b
Number of Factors
Cross Correlation

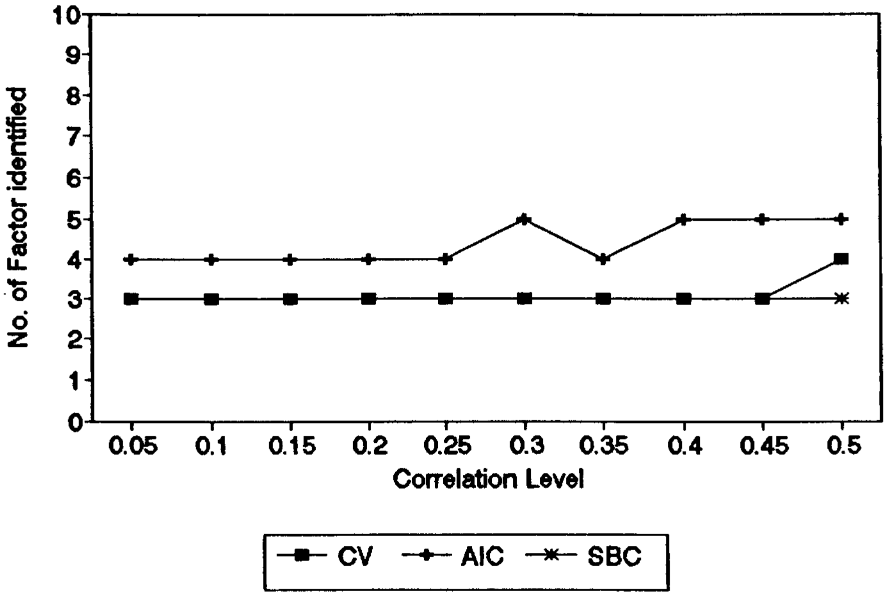

Figure 2.c
Number of Factors
Correlated Factor Structure