Approximate Factor Structures: Interpretations and Implications for Empirical Tests
Author(s): Mark Grinblatt and Sheridan Titman
Source: *The Journal of Finance*, Vol. 40, No. 5 (Dec., 1985), pp. 1367-1373
Published by: Blackwell Publishing for the American Finance Association
Stable URL: http://www.jstor.org/stable/2328118
Accessed: 26/11/2009 08:49

THE JOURNAL OF FINANCE • VOL. XL, NO. 5 • DECEMBER 1985

Approximate Factor Structures: Interpretations and Implications for Empirical Tests

MARK GRINBLATT and SHERIDAN TITMAN

ABSTRACT

This paper provides some new insights about approximate factor structures, as defined by Chamberlain and Rothschild [2], and their implications for empirical tests. First, we show that any economy that satisfies an approximate factor structure can be transformed, in a manner that does not alter the characteristics of investor portfolios, into an economy that satisfies an exact factor structure, as defined by Ross [9]. Second, we show that principal components analysis represents just one of many methods of forming groups of well-diversified portfolios with no idiosyncratic risk in large samples. Correct factor loadings will be obtained by regressing security returns on any group of these portfolios. Our interpretations of the Chamberlain and Rothschild results also provide additional insights into the testability of the Arbitrage Pricing Theory. We show that securities cannot be repackaged to hide factors in the manner suggested by Shanken [10] without the variance of some of the repackaged securities approaching infinity in large economies.

A RECENT PAPER BY Chamberlain and Rothschild [2] examined the implications of no arbitrage in a market with arbitrarily large numbers of assets. Their paper provides a number of interesting results, but perhaps its two main contributions are:

(i) the introduction of the concept of an approximate factor structure, which the authors argue is a weaker restriction on the asset return generating process than the perfect factor structure used in Ross' [9] development of the Arbitrage Pricing Theory, and,

(ii) the demonstration that under mild regularity conditions, principal components analysis provides asymptotically consistent (as the number of assets approaches infinity) estimates of the true factor loadings and is therefore superior to factor analysis because of its computational simplicity.

This paper uses standard portfolio theory to develop economic intuition for these two results. It then uses this intuition to reexamine their implications. Section I demonstrates that the notion of an approximate factor structure, while providing important insights, is not a significantly weaker restriction on the securities return generating process than Ross' exact factor structure. Any economy that satisfies the Chamberlain and Rothschild approximate factor

* Both authors are from the Graduate School of Management, University of California at Los Angeles. We wish to thank Philip Dybvig, Jonathan Ingersoll, Richard Roll, Bruce Rothschild, and Michael Rothschild for comments on an earlier draft of this paper.

structure can be transformed, in a manner that does not alter the characteristics of investor portfolios, into an economy that satisfies the Ross exact factor structure. This insight also applies to the Ingersoll [6] economy that yields the equal-weighted pricing bound, but not to the more general Ingersoll economy that yields a pricing bound on the weighted sum of squared deviations. Section II shows that principal components analysis, which forms portfolios to mimic factor returns, provides unbiased estimates of factor loadings in samples with infinite numbers of securities because it forms portfolios that are well-diversified (contain no idiosyncractic risk). This condition is practically nonrestrictive in large economies. Under the mild regularity conditions that make principal components a valid approach for APT tests, a continuum of groups of K well-diversified portfolios can be formed whenever an infinite number of asset returns are generated by a K-factor structure. Correct factor loadings for all assets in the economy can be recovered by regressing securities returns on any group of these portfolios. Hence, principal components analysis is just one of many possible techniques that can be used to test the APT. Our portfolio interpretation of the Chamberlain and Rothschild results also provides insights into Shanken's [10] critique of factor analysis tests of the APT. In Section III, we show that if securities are repackaged to hide factors, as Shanken suggests, then the variances of some of the securities in the economy approach infinity as the number of assets in the economy approaches infinity.

I. Factor Structure

In Ross' Arbitrage Pricing Theory, asset returns are generated by

$$\tilde{\underline{x}}_N = \underline{\mu}_N + \underline{\underline{B}}_N \tilde{\underline{f}} + \tilde{\underline{v}}_N$$

where $\tilde{\underline{x}}_N$ = random N − vector of ex-ante returns on the primitive assets

$\underline{\mu} = E(\tilde{\underline{x}}_N)$

$\underline{\underline{B}}_N$ = $N \times K$ matrix of factor loadings

$\tilde{\underline{f}}$ = random K − vector of factors, normalized to have zero mean, unit variance, and be uncorrelated with each other

$\tilde{\underline{v}}_N$ = random N − vector of mean-zero idiosyncratic disturbances, which are assumed to be uncorrelated with the factors and have diagonal covariance matrix $\underline{\underline{D}}_N$.

Letting $\underline{\underline{\Sigma}}_N = E[(\tilde{\underline{x}}_N - \underline{\mu}_N)(\tilde{\underline{x}}_N - \underline{\mu}_N)^T]$, these normalizations imply a decomposition of the covariance of returns into factor and nonfactor components. This decomposition is

$$\underline{\underline{\Sigma}}_N = \underline{\underline{B}}_N \underline{\underline{B}}_N^T + \underline{\underline{D}}_N.$$

Ross' APT examines a nested sequence of $\underline{\underline{\Sigma}}_N$, $\underline{\underline{B}}_N$, and $\underline{\underline{D}}_N$ matrices. His model essentially assumes that there is no arbitrage and that the diagonal elements of the nested sequence $\underline{\underline{D}}_N$ are uniformly bound above by $\zeta < \infty$ as N, the number of primitive assets, becomes infinite. This allows portfolios with weights of order

$1/N$ to asymptotically diversify away idiosyncractic risk in the limit. Indeed, as Chamberlain [1] has shown, any N-vector of portfolio weights (i.e., weights sum to one), $\underline{\alpha}_N$, satisfying $\underline{\alpha}_N^T \underline{\alpha}_N \to 0$ as $N \to \infty$ will completely eliminate idiosyncratic risk in the limit. To see this, note that

$$\underline{\alpha}_N^T \underline{\underline{D}}_N \underline{\alpha}_N \leq \zeta \underline{\alpha}_N^T \underline{\underline{I}}_N \underline{\alpha}_N = \zeta \underline{\alpha}_N^T \underline{\alpha}_N.$$

Since ζ is finite, $\underline{\alpha}_N^T \underline{\alpha}_N \to 0$ implies $\mathrm{var}(\underline{\alpha}_N^T \underline{\tilde{v}}_N) \to 0$.

It should be noted that a continuum of portfolios exist that diversify away idiosyncratic risk. To prevent arbitrage, all of these well-diversified portfolios must be asymptotically linearly priced according to their factor risk, (i.e., their factor loadings).[1] This is because some linear combination of K such portfolios with linearly independent loadings is perfectly correlated with any well-diversified portfolio. As others have formally shown (see Ross [9], Huberman [5], and Ingersoll [6]), this implies an approximate pricing relationship on the primitive assets:

$$\lim_{N \to \infty} (\underline{\mu}_N - \tau_0 \underline{l} - \underline{\underline{B}}_N \underline{\tau})^T (\underline{\mu}_N - \tau_0 \underline{l} - \underline{\underline{B}}_N \underline{\tau}) \leq \gamma < \infty, \quad \text{where } \underline{l} \text{ is a vector of ones.}$$

The work of Chamberlain and Rothschild (and Ingersoll) stresses that the idiosyncratic disturbances do not have to be uncorrelated for the APT to hold. In their definition of an approximate factor structure, a nondiagonal residual covariance matrix, $\underline{\underline{\Omega}}_N$, can replace $\underline{\underline{D}}_N$ in the previous analysis if the eigenvalues of $\underline{\underline{\Omega}}_N$ (which all must be positive) are uniformly bounded above by some finite number, λ. This corresponds to the K largest eigenvalues of $\underline{\underline{\Sigma}}_N$ becoming infinite as $N \to \infty$, with the remaining eigenvalues being uniformly bounded above if K is the dimension of the smallest factor structure that satisfies the above condition.

The analysis here demonstrates that any economy satisfying the Chamberlain and Rothschild conditions can be transformed to satisfy the Ross conditions, (and, conversely, Ross economies can be transformed into Chamberlain and Rothschild economies). The economies are transformed by repackaging the N security returns into N new returns constructed by forming N portfolios of the primitive assets. Investors are indifferent between the two economies, since, in choosing their portfolios, they can always "undo" the packaging of assets. The portfolio weights for repackaging are related to the eigenvectors of the covariance matrix of the idiosyncratic disturbances. By the definition of an eigenvector, if $\underline{\underline{S}}_N$ is an $N \times N$ matrix with columns equal to the eigenvectors of $\underline{\underline{\Omega}}_N$, then

$$\underline{\underline{S}}_N^{-1} \underline{\underline{\Omega}}_N \underline{\underline{S}}_N = \underline{\underline{\Lambda}}_N,$$

where $\underline{\underline{\Lambda}}_N$ is a diagonal matrix with diagonal elements that are the eigenvalues of $\underline{\underline{\Omega}}_N$. Because $\underline{\underline{\Omega}}_N$ is symmetric and positive definite, a well-known theorem from linear algebra states that the columns of $\underline{\underline{S}}_N$ are orthogonal and real, and the eigenvalues are real and positive. If we scale $\underline{\underline{S}}_N$, so that $\underline{\underline{S}}_N^T \underline{\underline{S}}_N = \underline{\underline{I}}_N$, then

$$\underline{\underline{S}}_N^T \underline{\underline{\Omega}}_N \underline{\underline{S}}_N = \underline{\underline{\Lambda}}_N.$$

[1] A rigorous proof of this is found in Ingersoll [6].

If the $\underline{\underline{S}}_N$ matrix had columns that were portfolio weights, the intuition behind the Chamberlain and Rothschild eigenvalue condition would become clear. In this case, $\underline{\underline{S}}_N^T \underline{\underline{\Omega}}_N \underline{\underline{S}}_N$ would be the covariance matrix of the idiosyncratic disturbances of a set of repackaged assets. The portfolios for repackaging would be the eigenvectors. The idiosyncratic disturbances in the repackaged asset market would be uncorrelated and, as in Ross' perfect factor structure, the variances of the idiosyncratic disturbances (which are the eigenvalues) would be uniformly bounded above. Since the APT pricing bound applies in the former asset market, it also holds in the equivalent asset market with correlated idiosyncratic risk.

There is, however, a minor technical difficulty with this argument. The columns of $\underline{\underline{S}}_N$ lie, not in the plane through the unit simplex, but in the unit (hyper)sphere, and thus, cannot be regarded as portfolio weights. This difficulty is overcome by rescaling the eigenvectors so that their elements sum to one. The rescaled (portfolio) matrix of eigenvectors $\underline{\underline{P}}_N$ is defined by

$$\underline{\underline{S}}_N = \underline{\underline{P}}_N \underline{\underline{\psi}}_N,$$

where $\underline{\underline{\psi}}_N$ is a diagonal matrix with diagonal entries that respectively form the vector $\underline{\underline{S}}_N^T \underline{l}$. If $\underline{\underline{P}}_N$ does not exist for the original Chamberlain-Rothschild economy, there is an intermediate economy with bounded eigenvalues (where assets are repackaged into portfolios) for which the diagonal elements of $\underline{\underline{\psi}}_N$ are not only nonzero, but uniformly bounded away from zero as $N \longrightarrow \infty$.[2] This demonstrates the following proposition.

PROPOSITION 1. *Every portfolio* $\underline{\alpha}_N$ *in the (Chamberlain-Rothschild) economy where there is correlated idiosyncratic risk, with covariances defined by* $\underline{\underline{\Omega}}_N$, *corresponds to a portfolio* $\underline{\underline{P}}_N^{-1}\underline{\alpha}_N$ *in the repackaged (Ross) economy where the idiosyncratic disturbances are uncorrelated, with covariances defined by the diagonal matrix* $\underline{\underline{D}}_N^* = (\underline{\underline{\psi}}_N^{-1})^2 \underline{\underline{\Lambda}}_N$. *Conversely, any portfolio* $\underline{\alpha}_N$ *in the Ross economy corresponds to a portfolio* $\underline{\underline{P}}_N \underline{\alpha}_N$ *in the Chamberlain and Rothschild economy. The elements of* $\underline{\underline{D}}_N^*$ *are uniformly bounded above if the eigenvalues of* $\underline{\underline{\Omega}}_N$ *are uniformly bounded above.*

Thus, the APT holds in the Chamberlain and Rothschild economy because it imposes essentially the same restrictions as Ross did in his derivation of the

[2] Consider the case where one or more of the eigenvectors of $\underline{\underline{\Omega}}_N$ have element sums that are not uniformly bounded away from zero. In this case, examine the class of all $N \times N$ portfolio matrices $\underline{\underline{Q}}_N$ with $\underline{\underline{Q}}_N^T 1 = 1$ and $\underline{\underline{Q}}_N^T \underline{\underline{Q}}_N$ having strictly positive uniformly bounded eigenvalues. $\hat{\underline{\underline{\Omega}}}_N = \underline{\underline{Q}}_N^T \underline{\underline{\Omega}}_N \underline{\underline{Q}}_N$ represents the covariance matrix of an economy that is repackaged with N linearly independent portfolios. The eigenvalues of $\hat{\underline{\underline{\Omega}}}_N$, given by $\hat{\underline{\underline{\Lambda}}}_N$, a diagonal matrix, are uniformly bounded over N, since the largest eigenvalue of $\underline{\underline{\Omega}}_N$ is less than or equal to the product of the largest eigenvalues of $\underline{\underline{\Omega}}_N$ and $\underline{\underline{Q}}_N^T \underline{\underline{Q}}_N$. There always exists a $\underline{\underline{Q}}_N$ such that the eigenvectors of $\hat{\underline{\underline{\Omega}}}_N$, defined by the columns of $\hat{\underline{\underline{S}}}_N$, have sums that are uniformly bounded away from zero. Denote $\hat{\underline{\underline{S}}}_N^* = \hat{\underline{\underline{S}}}_N \underline{\underline{\Psi}}_N^*$ to be the rescaled matrix of eigenvectors, where $\underline{\underline{\Psi}}_N^*$ is a diagonal matrix with elements that are uniformly bounded over N. It follows that

$$(\underline{\underline{Q}}_N \hat{\underline{\underline{S}}}_N^*)^T \underline{\underline{\Omega}}_N (\underline{\underline{Q}}_N \hat{\underline{\underline{S}}}_N^*) = \underline{\underline{\psi}}_N^{*2} \hat{\underline{\underline{\Lambda}}}_N.$$

APT. This result also applies to the Ingersoll [6] equally-weighted pricing bound, since the bounded variation condition he imposes is equivalent to the finite eigenvalue condition of Chamberlain and Rothschild. However, our analysis does not apply to the weighted sum of squares pricing bound in Ingersoll ([6], Equation (15)) since the correlation matrix in economies with correlated idiosyncratic residuals may not have a bounded norm, while the transformed economy with uncorrelated idiosyncratic residuals must always have a bounded norm. In this sense, the Ingersoll pricing result that weights the squared deviations from exact APT pricing is based on weaker economic assumptions than typical APT pricing bounds.

II. Diversification of Idiosyncratic Risk and Principal Components Analysis

Portfolios that diversify away all idiosyncratic risk are useful for empiricists.[3] Since they are perfectly correlated with the true factors, regressing security returns on the returns of these benchmark portfolios will yield unbiased estimates of the factor loadings and permit tests of the APT in large samples.[4] As we have already shown, portfolios $\underline{\alpha}_N$ with $\underline{\alpha}_N^T\underline{\alpha}_N \to 0$ in the Ross economy diversify away idiosyncratic risk. Hence, the analogous portfolio in the nondiagonal economy, $\underline{\underline{P}}_N\underline{\alpha}_N$, must also diversify away idiosyncratic risk in the latter economy. However, the next proposition, which is implicit in Chamberlain [1], demonstrates that any set of portfolio weights satisfying $\underline{\alpha}_N^T\underline{\alpha}_N \to 0$ also diversifies away idiosyncratic risk in the Chamberlain-Rothschild economy.

PROPOSITION 2. *Consider a sequence, indexed by N, of nonsingular nested positive definite residual covariance matrices, $\underline{\underline{\Omega}}_N$. Let λ_{1N} denote the largest eigenvalue of the matrix. A portfolio weighting sequence, $\underline{\alpha}_N$, diversifies away all idiosyncratic risk in the limit if*

(i) $\underline{\alpha}_N^T\underline{\alpha}_N \to 0$ as $N \to \infty$ and
(ii) $\sup_N\lambda_{1N} = \lambda < \infty$.

Proof: $\lambda\underline{\underline{I}}_N - \underline{\underline{\Lambda}}_N$ is positive semidefinite, where $\underline{\underline{\Lambda}}_N$ is the diagonal matrix of eigenvalues. Thus,

$$\underline{\alpha}_N^T\underline{\underline{\Omega}}_N\underline{\alpha}_N = \underline{\alpha}_N^T\underline{\underline{S}}_N\underline{\underline{\Lambda}}_N\underline{\underline{S}}_N^T\underline{\alpha}_N \leq \lambda\underline{\alpha}_N^T\underline{\underline{S}}_N\underline{\underline{S}}_N^T\,\underline{\alpha}_N = \lambda\underline{\alpha}_N^T\underline{\alpha}_N.$$

Thus if, $\underline{\alpha}_N^T\underline{\alpha}_N \to 0$, $\lambda\underline{\alpha}_N^T\underline{\alpha}_N \to 0$, implying $\underline{\alpha}_N^T\underline{\underline{\Omega}}_N\underline{\alpha}_N \to 0$. Q.E.D.

[3] Such portfolios generally require an infinite number of assets. With a finite number of assets, portfolios that minimize idiosyncratic risk will yield consistent estimates of the factor loadings. The portfolio weights of these ideal benchmarks are proportional to $\underline{\underline{\Omega}}^{-1}\underline{\underline{B}}(\underline{\underline{B}}^T\underline{\underline{\Omega}}^{-1}\underline{\underline{B}})$. In practice, however, there is either great computational complexity in obtaining these portfolios or an identification problem in estimating both $\underline{\underline{\Omega}}$ and $\underline{\underline{B}}$. Factor analysis assumes that $\underline{\underline{\Omega}}$ is diagonal; principal components analysis assumes that $\underline{\underline{\Omega}}$ is an identity matrix. For a further discussion, see Lehmann and Modest [8] or Grinblatt and Titman [4].

[4] See Jobson [7].

This explains why principal components analysis yields unbiased factor loadings. Any weighting scheme with the squared sum of the weights converging to zero yields asymptotically correct estimates. Given a particular factor structure, the asymptotic principal components benchmark portfolio weights are proportional to the columns of the matrix $\underline{\underline{B}}_N(\underline{\underline{B}}_N^T\underline{\underline{B}}_N)^{-1}$ as $N \rightarrow \infty$. As long as the true factor loadings don't "explode" (i.e., are bounded above in absolute value) and there are an infinite number of securities with nonzero factor loadings for each factor, principal components chooses weights of order $1/N$, which result in well-diversified portfolios. In the principal components orthogonalization and rotation of the estimated factors, the j^{th} column of $\underline{\underline{B}}_N$, $j = 1, \cdots, K$, is estimated to be

$$\theta_{jN}^{1/2}\underline{t}_{jN},$$

where θ_{jN} is the j^{th} largest eigenvalue of $\underline{\underline{\Sigma}}_N$, and $\underline{t}_{jN}$ is the associated eigenvector on the unit sphere. Thus, the asymptotic squared sum of the portfolio weights for the factor j proxy is the j^{th} diagonal element of the normalized diagonal matrix $(\underline{\underline{B}}_N^T\underline{\underline{B}}_N)^{-1}$, which is $\lim_{N\rightarrow\infty} 1/\theta_{jN}$. By assumption, the K largest eigenvalues are unbounded, making this expression zero.

The unboundedness of the K largest eigenvalues implies that the dimension of the subspace occupied by the (nonzero-beta) well-diversified portfolio returns equals the number of factors. This need not necessarily be the case. Suppose, for instance, that three factors affect an infinite number of assets, but a fourth factor affects only a finite number of assets. A technique like principal components cannot yield more than three well-diversified portfolios. A variety of portfolios of the finite number of assets sensitive to the fourth factor will serve as adequate proxies (including the fourth principal components portfolio), each yielding a different set of factors loadings and market risk premiums. We avoid nonuniqueness only by imposing the unboundedness condition.[5]

III. The Shanken Critique

The portfolio interpretation of the Chamberlain and Rothschild [2] results in the previous two sections also provides insights into Shanken's [10] critique of factor analysis. This critique is based on the notion that the packaging of primitive securities might hide factors. Although Shanken does not explicitly restrict his analysis to small subsets of the economy, the Chamberlain and Rothschild analysis implies that the degeneracies he describes in his paper cannot occur in large samples. Chamberlain and Rothschild have shown that if K common factors generate returns, then the K largest eigenvalues of the covariance matrix $\underline{\underline{\Sigma}}_N$ grow without bound as N gets large. This implies that if securities are repackaged to eliminate m factors, the variance of m of the new securities will approach infinity as the number of securities in the economy approaches infinity. Consider, for example, the case where the economy is repackaged so that securities are all uncorrelated. Since the variances of these repackaged securities are of the order of the eigenvalues of the original covariance matrix, it follows

[5] Ingersoll ([6], Equation (30)) imposes a similar regularity condition by requiring that the smallest eigenvalue of $\underline{\underline{B}}_N^T\underline{\underline{B}}_N$ be unbounded.

that the K largest variances are infinite. These repackaged securities will be (correctly) selected by maximum likelihood factor analysis as common factors; hence, in sufficiently large samples, factor analysis (among other techniques, including principal components analysis) will select the correct number of factors, and the estimated factor loadings will be consistent estimates of the factor sensitivities.

IV. Conclusions

This paper has shown that an approximate factor structure is not significantly less restrictive than the exact factor structure proposed by Ross. This does not imply that the notion of an approximate factor structure provides no important insights about the APT. For example, it tells us that the existence of intraindustry residual correlations will not necessarily violate the assumptions of the APT. As long as the N primitive securities can be combined into N portfolios that satisfy the Ross conditions, the APT will hold. The Chamberlain and Rothschild [2] analysis also provides important insights about the Shanken [10] critique. An economy can be repackaged to have a factor structure of lower dimension only by creating new securities with infinite variances. However, any statistical technique that forms well-diversified portfolios, such as principal components analysis and maximum likelihood factor analysis, will correctly interpret these infinite variance securities as factors.

We have also shown that any technique that forms benchmark portfolios with weights of order $1/N$ will provide unbiased estimates of the factor loadings for infinite N. This explains why principal components analysis "works." It also suggests that other techniques that make use of this diversification principle (e.g., Chen [3]) will yield unbiased estimates if portfolios are formed from infinite samples of securities.

REFERENCES

1. G. Chamberlain. "Funds, Factors, and Diversification in Arbitrage Pricing Models." *Econometrica* 51 (September 1983), 1305–23.
2. G. Chamberlain, and M. Rothschild. "Arbitrage, Factor Structure, and Mean-Variance Analysis on Large Asset Markets." *Econometrica* 51 (September 1983), 1281–304.
3. N. Chen. "Some Empirical Tests of the Theory of Arbitrage Pricing." *Journal of Finance* 38 (December 1983), 1393–414.
4. M. Grinblatt, and S. Titman. "The Relation between Mean-Variance Efficiency and Arbitrage Pricing." UCLA Working Paper 4-83, 1985.
5. G. Huberman. "A Simple Approach to Arbitrage Pricing Theory." *Journal of Economic Theory* 28 (October 1982), 183–91.
6. J. Ingersoll. "Some Results in the Theory of Arbitrage Pricing." *Journal of Finance* 39 (September 1984), 1021–39.
7. J. Jobson. "A Multivariate Linear Regression Test for the Arbitrage Pricing Theory." *Journal of Finance* 37 (September 1982), 1037–42.
8. B. Lehmann and D. Modest. "The Empirical Foundation of the Arbitrage Pricing Theory II: The Optimal Construction of Basis Portfolios." Columbia University Working Paper, 1985.
9. S. Ross. "The Arbitrage Theory of Capital Asset Pricing." *Journal of Economic Theory* 13 (December 1976), 341–60.
10. J. Shanken. "The Arbitrage Pricing Theory: Is it Testable?" *Journal of Finance* 37 (December 1982), 1129–40.