A Test of APT with Maximum Sharpe Ratio

CHU ZHANG*

This version: January, 2006

*Department of Finance, The Hong Kong University of Science and Technology, Clear Water Bay, Kowloon, Hong Kong (phone: 852-2358-7684, fax: 852-2358-1749, e-mail: czhang@ust.hk). I would like to thank Peter Xiang for computational assistance, Raymond Kan, Mungo Wilson, and Guofu Zhou for helpful comments on an early version of the paper. Financial support from the RGC Competitive Earmarked Research Grant HKUST6268/04H is gratefully acknowledged. All remaining errors are mine.

A Test of APT with Maximum Sharpe Ratio

ABSTRACT

This paper tests the asymptotic arbitrage pricing theory (APT) on individual stocks with the factors extracted by the Connor-Korajczyk method. The asymptotic APT fails if and only if the number of unbounded eigenvalues of the second-moment matrix of excess returns exceeds that of the variance matrix of excess returns by one. A test is developed using this theoretical result based on the eigenvectors extracted using the CK method. The test statistic is shown to be related to the maximum Sharpe ratio among portfolios of all individual stocks. The empirical evidence, supplemented by simulation results, lends support to the implication of the asymptotic arbitrage pricing theory.

Mainstream asset pricing theory is based on the notion of systematic risks represented by marketwide factors. Expected returns on individual securities are linear functions of their standardized covariances, or betas, with the factors. The Capital Asset Pricing Model (CAPM) developed by Sharpe (1964) is the first example of a beta pricing model in which the return on the market portfolio is the single factor that matters for determining expected returns on individual securities. The Inter-temporal Capital Asset Pricing Models (ICAPM) of Merton (1973) and the Arbitrage Pricing Theory (APT) of Ross (1976) extend the CAPM to multi-factor settings. Despite their vintage, these models remain the most commonly used in theoretical and empirical analysis by academics and practitioners.

The most popular way of testing a beta pricing model has been the method of Gibbons, Ross and Shanken (1985), which examines a measure of deviations from the beta pricing model for certain given portfolio returns as mimicking portfolios for systematic factors. The test basically examines whether a given set of portfolio returns is a sufficient set of factors. If the test rejects a particular model, it is always possible that the beta-pricing principle is correct, while the factors in the model are mis-specified. In addition, the test requires that the number of securities be smaller than the number of time-series observations in the return data used for the test. As a result, the beta pricing model is typically tested on a small number of portfolios sorted by certain stock characteristics.

On another front, methods of extracting factors from individual stock returns have been developed in the factor structure of the APT. The asymptotic APT extended by Chamberlain and Rothschild (1983) relaxes the strict factor assumption made by Ross (1976) and others. Connor and Korajczyk (CK, 1986, 1988) develop a method to deal with the case of a large number of securities and a much smaller number of time-series observations. The CK method has been widely used to extract factors from individual stock returns and to test asset pricing anomalies. For example, McCulloch and Rossi

(1991) use the CK method to test the firm size anomaly and Brennan et al. (1998) extract the CK factors to examine the size and book-to-market effects. Surprisingly, however, the implication of the asymptotic APT that errors from the beta pricing should be contained has never been tested although the method of extracting factors has been well developed. In fact, there is a debate in the literature about whether the APT implication can be meaningfully tested. Shanken (1992) summarizes the controversy surrounding the testability issue.

The purpose of the current paper is to develop a test of the beta pricing implication of the asymptotic APT using factors extracted by the CK method from individual stock returns. Despite the difficulty pointed out by Shanken, the basic implication of the asymptotic APT can be tested. More importantly, the implication of the asymptotic APT should be tested before extracted factors are used to examine asset pricing anomalies because, if the implication of the asymptotic APT is invalid, the "factors" extracted by any method will be contaminated by pricing errors, so the result of examining asset pricing anomalies using these "factors" may not be relied upon.

The simple test presented in this paper is based on a theoretical result that, if the implication of the asymptotic APT does not hold, the number of unbounded eigenvalues from the (non-centered) second-moment matrix of the stock returns equals one plus the number of unbounded eigenvalues of the variance matrix of the stock returns, as the number of testing securities goes to infinity. Theoretically speaking, by examining the number of unbounded eigenvalues of the second-moment matrix and of the variance matrix, the implication of the asymptotic APT can be tested. In practice, there are only a finite number of testing securities, so determining the number of unbounded eigenvalues is tricky. The literature on determining the number of factors in individual US stock returns has not reached a consensus. To bypass this difficulty, the current paper uses a test conditioned on the given number of extracted eigenvectors using the CK method. If the implication of the asymptotic APT is violated, for a certain number

of extracted factors, the test should detect the violation. On the other hand, if the test does not detect any violation of the APT implication for all potentially reasonable numbers of factors, then it is safe to conclude that the implication of the asymptotic APT is upheld by the empirical evidence. The asymptotic distribution as the number of testing securities tends to infinity is derived for the test statistic.

The asymptotic APT is intimately related to the notion of the (generalized) Sharpe ratio, defined as the expected payoff of a zero-cost investment divided by its standard deviation. An asymptotic arbitrage opportunity is defined as a situation in which a sequence of portfolios can be formed such that their Sharpe ratios tend to infinity as the number of stocks in the portfolios increases to infinity. Chamberlain and Rothschild (1983) establish that the squared norm of the pricing error from the beta pricing equation is bounded by the product of the maximum squared Sharpe ratio and the maximum variance of idiosyncratic risk. The test statistic developed in this paper, based on the minimum eigenvalue of the variance matrix of the extracted eigenvectors, turns out to be related to the sample version of the maximum squared Sharpe ratio of the portfolios formed from all individual stocks. This provides a nice interpretation of the proposed test statistic.

The method of testing the asymptotic APT is applied to data from US individual stocks traded on NYSE/AMEX/NASDAQ that have complete observations during eight non-overlapping sixty-month sample periods between 1965 and 2004. For all the eight sample periods, the test does not find evidence against the implication of the asymptotic APT. Although the maximum squared sample Sharpe ratios tend to be much higher than those of the value-weighted and equally weighted market portfolio of all the NYSE/AMEX/NASDAQ stocks, they are well contained in the region that does not violate the implication of the asymptotic APT. Since the inference is based on asymptotic distribution of the test statistic, a simulation exercise is carried out to ensure that the results based on the finite sample of actual data are not due to inferential error.

The rest of this paper is organized as follows. Section 1 presents theoretical results on the implication of the asymptotic APT regarding the number of unbounded eigenvalues of the second-moment and variance matrices. The connection between extracted factors and the Sharpe ratio is given there as well, followed by the test statistic based on the minimum eigenvalue of the variance matrix of the extracted eigenvectors. Section 2 reports empirical results on US individual stock returns from eight sixty-month sample periods. Section 3 conducts simulations confirming the results from actual data. The last section concludes the paper and an appendix contains the proof of the analytical results.

I. Methodology

A. Theoretical Foundation

Suppose r_t is the vector of returns in excess of the riskfree rate on n securities in month t. In a factor model, the excess returns are driven by

$$r_t = a + Bf_t + \varepsilon_t, \tag{1}$$

where a is an n-vector of constants, known as Jensen's alpha, B is an $n \times k$ matrix of betas of, or loadings on, the k-vector of factors, f_t, at t, and ε_t is the idiosyncratic risk of the n securities, satisfying $E(\varepsilon_t|f_t) = 0$. The vector a is the focus of this paper. Without loss of generality, we can assume that a is orthogonal to B, i.e., $a'B = 0$. Otherwise, we can let $\tilde{a} = a - B(B'B)^{-1}B'a$, $\tilde{f}_t = f_t + (B'B)^{-1}B'a$ and rewrite (1) as $r_t = \tilde{a} + B\tilde{f}_t + \varepsilon_t$. The factors are unobserved with mean μ_f and variance Σ_f. Without loss of generality, Σ_f is assumed to be positive definite, which means that f_t contains no redundancy. The idiosyncratic risk has a mean of zero and a variance matrix denoted Σ_ε. The variance matrix of r_t is $\Sigma_r = B\Sigma_f B' + \Sigma_\varepsilon$.

In the strict factor model, Σ_ε is assumed to be a diagonal matrix, so the components of ε_t are uncorrelated. Chamberlain and Rothschild (1983) relax the strong assumption

of the uncorrelated idiosyncratic risk and propose an approximate factor model. The approximate factor model is characterized by the structure of the variance matrix of r_t, Σ_r. As the number of securities, n, grows, if k largest eigenvalues of Σ_r increase without bound while the other eigenvalues remain bounded, the returns are said to be generated from an approximate k-factor model. In the approximate factor model, the idiosyncratic risk can be cross-sectionally correlated. The implication of the asymptotic APT is that the squared norm of the pricing error, $a'a$, is bounded. If the asymptotic APT holds, the returns in an approximate factor model can also be written as (1), where the factors can be taken as the k linear combinations of the returns where the coefficients are the k eigenvectors of Σ_r corresponding to the k largest eigenvalues and the idiosyncratic risks are the linear combinations of returns associated with the remaining eigenvectors.

If the implication of the asymptotic APT is violated, i.e., $a'a$ is unbounded, we say that a is a systematic pricing error. Otherwise, when $a'a$ is bounded as the asymptotic APT claims, we say that the pricing error is unsystematic. The notion of a systematic pricing error parallels the notion of systematic factors with unbounded eigenvalues. This can be seen more easily in a one-factor model, $r_t = bf_t + \varepsilon_t$, with $\sigma_f = 1$ and uncorrelated idiosyncratic risks. The second-moment matrix of the returns has one unbounded eigenvalue equal to $b'b$.

When extracting factors from the variance matrix of the return, the effect of a is absent because the variance is invariant to a mean shift. The effect of a nonzero a is, however, retained in the second-moment matrix. Let $S_r = E(r_t r_t')$ be the second-moment matrix of r_t. From (1),

$$S_r = aa' + BS_fB' + \Sigma_\varepsilon, \tag{2}$$

where $S_f = Ef_t f_t'$ is the second-moment matrix of f_t. A test of existence of a systematic pricing error can be based on the following proposition.

Proposition 1. Suppose the returns follow an approximate k-factor model. Let k^* be the number of unbounded eigenvalues of the second-moment matrix of excess returns,

S_r. Then, as n goes to infinity, $k^* = k + 1$ if $a'a$ is unbounded, or $k^* = k$ if $a'a$ is bounded.

The proposition gives a simple criterion that can be used to test the existence of a systematic pricing error. If the pricing error is unsystematic, then the number of unbounded eigenvalues of the second-moment matrix of excess returns equals that of the variance matrix. If the pricing error is systematic, then the number of unbounded eigenvalues of the second moment matrix of excess returns equals that of the variance matrix plus one.

By writing $a + Bf_t = (a \ B)(1 \ f_t')'$, we see that the difference between a k-factor model with a pricing error and a $(k+1)$-factor model without a pricing error is that, in the model with a pricing error, the "factors" contain a constant. Since factors are not observed, any inference about pricing error must be based on extracted factors. Suppose that g_t is a k^*-vector of extracted "factors" from the second-moment matrix of returns. From Proposition 1, we know that, in the limit as the number of stocks tends to infinity, g_t is a linear transformation of f_t if there is no systematic pricing error, or a linear combination of $(1 \ f_t')'$ if there is a systematic pricing error. We can write $g_t = Cf_t^*$ where the dimension of g_t is k^*. Either $f_t^* = f_t$ or $f_t^* = (1 \ f_t')'$ depending on whether there is a systematic pricing error. The matrix C is nonsingular because Σ_f is assumed to be positive definite. In terms of extracted "factors", the difference between a k-factor model with a pricing error and a $(k+1)$-factor model without a pricing error is that, in the former situation, the variance matrix of g_t is degenerate, while in the latter, the variance matrix is positive definite.

In the empirical work in the next section, we use extracted factors based on the Connor-Korajczyk method. Extracted factors are linear combinations of returns. As Chamberlain and Rothschild (1983) show, extracted factors can be linear combinations of the factors themselves, if the variance matrix of the return is known. In particular, the Connor-Korajczyk factors are normalized because they are orthonormal eigenvectors

of the sample second-moment matrix. Suppose g_t is normalized to have $Eg_tg_t' = I_k$ and the mean of g_t is denoted as μ_g. The proposition below characterizes the theoretical properties of the normalized "factors" and relates them to the maximum squared Sharpe ratio of portfolios of the excess returns.

Proposition 2. Let g_t be the k^*-vector of "factors" extracted from the second-moment matrix of the return, normalized to have $Eg_tg_t' = I_{k^*}$. Denote $\mu_g = Eg_t$ and $\gamma = \mu_g'\mu_g$. Then, (i) $0 \leq \gamma < 1$. (ii) There is a systematic pricing error if and only if $\gamma = 1$. (iii) The maximum squared Sharpe ratio equals

$$s = \frac{\gamma}{1-\gamma} \tag{3}$$

or infinity if $\gamma = 1$.

Before we move to the econometric method and the empirical analysis, some remarks about the testability of the APT are in order. The APT is proposed as an alternative to the well-known CAPM. However, the testability of the APT has long been debated in the literature. The version of APT derived by Chamberlain and Rothschild (1983) states that

$$a'a \leq s\bar{\sigma}_\varepsilon, \tag{4}$$

where $\bar{\sigma}_\varepsilon$ is an upper bound of the idiosyncratic risk. Since the standard APT does not make assumptions on investors' preference, the theory does not involve the determination of the maximum squared Sharpe ratio, s. If the implication of the APT is confined to the inequality (4) with s unspecified, then, as Shanken (1992) correctly points out, testing the APT is not very meaningful because that inequality constraint is more or less a tautology. An approximate linear relationship between the expected excess return and the betas is also difficult to test. However, the spirit of the APT is beyond the inequality. For the theory to be useful, so that expected returns can be well approximated by linear functions of factor betas, the maximum squared Sharpe ratio has to be small, despite the notion of no asymptotic arbitrage only rules out the case in which s equals infinity. What

data say about the existence of asymptotic arbitrage opportunities and, more generally, about the value of the maximum squared Sharpe ratio is certainly a meaningful question to ask. For practical purposes, a large s can be regarded as the failure of the asymptotic APT although how large is large is quite subjective. As will be shown later, the sample squared Sharpe ratio of the value-weighted market portfolio in the US market during the 1965-2004 period is less than 0.01. A conservative choice for testing the maximum squared Sharpe ratio can set, say, $s \geq 9$. If such a choice is taken, then one rejects the asymptotic APT only if there is evidence that shows s is greater than 9. A more radical choice of s could be, say, $s = 1$. In that case, as long as evidence points to $s \geq 1$, a much less stringent condition than $s \geq 9$, one refutes the asymptotic APT.

Not only could the implication of the asymptotic APT be tested and meaningfully interpreted, it also should be tested before extracted "factors" are used to examine asset pricing anomalies as if these "factors" are true factors. To illustrate the point, suppose the beta pricing model does not hold and the excess returns are generated by

$$r_t = a + bf_t + \varepsilon_t$$

where a is a systematic pricing error with $a'b = 0$ and f_t is the return on the market portfolio with positive variance, the only factor of the excess returns. Let x be a constant n-vector of a firm-specific variable that is highly, cross-sectionally correlated with a. Naturally the expected excess returns are found to be related to the market beta, b, and the firm-specific variable, x, because x is highly correlated with a. To see whether x is a proxy for the beta of an unobserved factor, one extracts "factors" from the returns and obtains $g_t = C(1 \ f_t)'$ without any estimation error where C is a 2×2 nonsingular matrix. The $n \times 2$ beta matrix of the extracted "factors" can be worked out as

$$B_g = E(r_t g_t')[E(g_t g_t')]^{-1} = (a \ \ b)C^{-1},$$

so B_g spans the same subspace as $(a \ \ b)$. Given B_g, x will have no additional explanatory power for the expected excess returns. One then claims that the puzzle of dependence

of expected returns on x is solved and a two-factor beta-pricing model holds. That inference, of course, is erroneous. This highly simplified example illustrates the danger of using extracted "factors" without verifying their validity.

B. Econometric Method

Proposition 1 suggests that we can test the asymptotic APT by examining the number of unbounded eigenvalues in Σ_r and S_r. However, the econometric issues involved in determining the numbers from the *sample version* of Σ_r and S_r turn out to be difficult to solve. Trzcinka (1986) find that the largest eigenvalue of the variance matrix dominates the rest of the eigenvalues. Brown (1989), however, argues convincingly that one dominant eigenvalue does not mean that there is only one systematic factor in the returns. Connor and Korajczyk (1993) propose a method to determine the number of factors that measures the marginal contribution of an additional factor and report findings of one to six factors for various subperiods. Geweke and Zhou (1996) apply a Bayesian approach and find little improvement in reducing pricing errors by having additional factors beyond the first one. Bai and Ng (2001) design various test statistics under general factor structures and draw the conclusion that two factors are adequate for the US stock returns. In short, there has been no consensus in the literature on what is the best way to determine the number of systematic factors and there has been no consensus on the actual number of factors found in the US stock market.

In this paper, we explore a method that conditions on any empirically relevant number of factors. The method first follows Connor and Korajczyk (1986, 1988) to extract possibly contaminated factors and then tests if these factors contain systematic pricing errors. It is based on the minimum eigenvalue of the variance matrix of the extracted "factor," which is also related to the maximum squared Sharpe ratio. Let $R = (r_1, \cdots, r_\tau)$ be the observations of r_t for $t = 1, \cdots, \tau$, $F = (f_1, \cdots, f_\tau)$, and

$\varepsilon = (\varepsilon_1, \cdots, \varepsilon_\tau)$. In matrix form,

$$R = a1'_\tau + BF + \varepsilon, \tag{5}$$

where 1_τ is the τ-vector of ones. Let

$$\Omega = \frac{1}{n} R'R. \tag{6}$$

For a given number k^* as a candidate of the number of unbounded factors, the Connor and Korajczyk (1986, 1988) method is to extract (transformed) factors as the $k^* \times \tau$ matrix of eigenvectors corresponding to the largest k^* eigenvalues of Ω. It is shown in the case of $a = 0$ that the eigenvectors converge to a nonsingular linear transformation of the factors, F, as n goes to infinity. They do not consider the pricing errors, however. With a systematic pricing error, a, the extracted factors are a transform of $(1_\tau \ F')'$ in the limit as n goes to infinity, rather than just F, as we can see formally that $R = (a \ B)(1_\tau \ F')' + \varepsilon$.

As in the case of Σ_r versus S_r, the existence of a systematic pricing error will show up in the difference between the second-moment and variance matrices of the eigenvectors. Let $\tau^{-1/2}G$ be the $k^* \times \tau$ orthonormal matrix of the k^* eigenvectors of Ω corresponding to the k^* largest eigenvalues of Ω, arranged as row vectors. Let $\bar{g} = \frac{1}{\tau} G1_\tau$. By construction, $\tau^{-1}GG'$ has k^* eigenvalues all equal to one. To see whether all the k^* eigenvectors are true factors or one of them is a pricing error after transformation, we can examine the smallest eigenvalue of $\tau^{-1}(G - \bar{g}1'_\tau)(G - \bar{g}1'_\tau)'$, denoted as λ_{k^*}. If the pricing error is systematic, then in the limit as n goes to infinity, $\tau^{-1/2}G$ is transformed from $(1_\tau \ F')$, which contains a constant row. As a result, $\tau^{-1}(G - \bar{g}1'_\tau)(G - \bar{g}1'_\tau)'$ will be degenerate. This can be found by testing $\lambda_{k^*} = 0$. The conditions and the distribution of the formal test is stated in the following proposition.

Proposition 3. Suppose $\Omega = \frac{1}{n} R'R$ is a $\tau \times \tau$ positive definite random matrix, the $k^* \times \tau$ matrix $\tau^{-1/2}G$ is the k^* orthonormal eigenvectors of Ω corresponding to the largest k^* eigenvalues, $\bar{g} = \frac{1}{\tau} G1_\tau$ and $\bar{\gamma} = \bar{g}'\bar{g}$. Then,

(i) $\tau^{-1}(G - \bar{g}1'_\tau)(G - \bar{g}1'_\tau)'$ has $k^* - 1$ eigenvalues equal to one and one eigenvalue, $\lambda_{k^*} = 1 - \bar{\gamma}$, between zero and one, both inclusive.

(ii) Suppose F is normally distributed. Then, the asymptotic distribution of $\bar{\gamma}$ is a non-central Beta with density function

$$p(x) = \frac{\Gamma(\frac{\tau}{2})}{\Gamma(\frac{\tau-k^*}{2})\Gamma(\frac{k^*}{2})} x^{\frac{k^*}{2}-1}(1-x)^{\frac{\tau-k^*}{2}-1} e^{\frac{-\tau s}{2}} {}_1F_1\left(\frac{\tau}{2}, \frac{k^*}{2}, \frac{\tau s x}{2}\right), \quad 0 \leq x \leq 1, \quad (7)$$

where $s = \gamma/(1-\gamma)$ is the maximal squared Sharpe ratio, $\gamma = \mu'_g \mu_g$, $\mu_g = Eg_t$, and ${}_1F_1(\cdot,\cdot,\cdot)$ is the confluent hypergeometric function.

When $s = 0$, the distribution is the familiar central Beta distribution with the degrees of freedom $(\frac{k^*}{2}, \frac{\tau-k^*}{2})$. In typical applications, $\tau \gg k^*$, so the majority of the mass of the distribution leans toward zero. However, the case of $s = 0$, corresponding to the case of risk-neutral investors, is not interesting. A non-central Beta distribution with a non-centrality parameter, $\tau s/2$, shifts the mass of the distribution to the right as s becomes greater. In the extreme case where $s \to \infty$, the distribution becomes degenerate and concentrates on one. Figure 1 below depicts three such non-central beta density functions corresponding to $s = 0.25, s = 1$, and $s = 9$ (or $\gamma = 0.2$, 0.5, and 0.9) with $\tau = 60$ and $k^* = 1, 3, 5, 10$. One observation from the figure is that, as k^* becomes greater, the distribution shifts more to the right. For $k^* = 1$, the modal point of the distribution is close to the parameter γ. For $k^* = 10$, however, the modal point of the distribution is much greater than the parameter γ.

Figure 1 here

Proposition 3 suggests the following way of testing the asymptotic APT. For a plausible value of k^*, we take k^* orthonormal eigenvectors of Ω, $\tau^{-1/2}G$. We then calculate the smallest eigenvalue of $\tau^{-1}(G - \bar{g}1'_\tau)(G - \bar{g}1'_\tau)'$, $\hat{\lambda}_{k^*}$. If $\hat{\lambda}_{k^*}$ is close to zero, then we can reject the asymptotic APT. On the other hand, if $\hat{\lambda}_{k^*}$ is close to one, then we do not

reject the asymptotic APT. Equivalently, we can calculate $\bar{\gamma} = 1 - \hat{\lambda}_{k^*}$ from $\bar{g}$ directly. we reject the asymptotic APT if $\bar{\gamma}$ is close to one. The rejection decision can be based on the asymptotic distribution in (7). In particular, if we set the null hypothesis to $s = \infty$, i.e., the APT fails, we can test the hypothesis by calculating the left-tail p-value of the test statistics, $\bar{\gamma}$ or $\bar{s}$, using the distribution corresponding to a large γ (or s) value, say $\gamma = 0.9$ (or $s = 9$). Such a p-value is an upper bound of the true upper bound. If the hypothesis $\gamma \geq 0.9$ (or $s \geq 9$) is rejected, the evidence is then in favor of the implication of the asymptotic APT.

II. Empirical Results

The data used for empirical tests are the monthly stock returns at the firm level in the US between 1965 and 2004. The data are from the Center for Research in Security Prices at the University of Chicago. All the common stock traded on NYSE/AMEX/NASDAQ are included except for American Deposit Receipts. More specifically, preferred stocks are excluded, but common stocks of utility companies and financial companies are not excluded.

Following the convention in the empirical asset pricing literature, the entire sample period is broken into eight 60-month subperiods and tests of the asymptotic APT are conducted within each subperiod. For each subperiod, stocks with non-missing monthly returns are collected and the second moment matrix, $\tilde{\Omega} = \frac{1}{\tau} R'R$, is calculated where R is the $n \times \tau$ excess returns with n being the number of stocks with non-missing returns and $\tau = 60$ being the number of observations for each stock. Note that $\tilde{\Omega}$ differs from Ω by a factor of n/τ and $\tilde{\Omega}$ has the same eigenvalues as the second-moment matrix of the excess returns, $\frac{1}{n} RR'$. From Ω, 60 eigenvalues are obtained and are arranged in descending order. Table 1 reports for each subperiod the first ten eigenvalues, $l_j,\ j = 1, \cdots, 10$, the proportion of each eigenvalue in the total, l_j/L, where $L = \sum_{i=1}^{60} l_i$, the first j cumulative eigenvalues, $\sum_{i=1}^{j} l_i$, and the proportional first j cumulative eigenvalues, $\sum_{i=1}^{j} l_i/L$.

Table 1 here

The number of stocks in each period ranges from 1487 to 3858 and the sum of the eigenvalues, L, ranges from 21 to 139. Early periods have relatively smaller numbers of stocks with non-missing values. The largest sum of the eigenvalues occurs in the last subperiod in which stock prices are very volatile. The big differences across subperiods is one reason for testing within each subperiod. The first eigenvalue accounts for 12% to 34% of the total sum of all the eigenvalues. The second eigenvalue is considerably smaller than the first one. The rest of them decline gradually without showing a clear point on the number of dominant eigenvalues. Though not reported here, the pattern continues until $j = 60$. This phenomenon has been documented in the literature and causes difficulty in determining the number of factors in the stock returns.

The Connor-Korajczyk method is then used to extract factors from the second moment matrix of the excess returns, $\tilde{\Omega}$. For an integer k^*, let $\tau^{-1/2}G = \tau^{-1/2}(g_1, \cdots, g_\tau)$ be the k^* eigenvectors of Ω corresponding to the largest k^* eigenvalues of $\tilde{\Omega}$. As explained earlier, the smallest eigenvalue of $\tau^{-1}(G - \bar{g}1'_\tau)(G - \bar{g}1'_\tau)'$ is $1 - \bar{\gamma}$ where $\bar{\gamma} = \bar{g}'\bar{g}$ and $\bar{g}$ is the k^*-vector mean of G. Proposition 3 establishes that $\bar{\gamma}$ has a noncentral Beta distribution whose non-centrality parameter is a multiple of s, where s is the asymptotic maximum squared Sharpe ratio. Table 2 reports the statistic $\bar{\gamma}$ for $k^* = 1, 2, \cdots, 10$. Beside $\bar{\gamma}$ is $\bar{s} = \bar{\gamma}/(1 - \bar{\gamma})$. For the calculated statistic $\bar{\gamma}$, its left-tail p-values under the hypothesis $\gamma = 0.9$, $\gamma = 0.5$, and $\gamma = 0.2$, which correspond to $s = 9$, $s = 1$, and $s = 0.25$, are reported.

Table 2 here

The test statistic $\bar{\gamma}$ ranges from zero to 0.20, depending on the subperiod and k^*s. The left-tail p-values at $\gamma = 0.9$ and $\gamma = 0.5$ are virtually zero and the left-tail p-value at $\gamma = 0.2$ is also very low and only occasionally goes above 0.10. A low value of γ means that there does not appear to be any unconditional portfolio strategy that can generate

a high enough Sharpe ratio to be termed as an asymptotic arbitrage opportunity under any reasonable subjective definition of an arbitrage opportunity. As far as the number of possible factors considered here, evidence is strong to uphold the implication of the asymptotic APT. For k^* greater than 10, although $\bar{\gamma}$ and $\bar{s}$ increase, the left-tail p-values do not necessarily increase, as we see from the pattern in $p_{0.2}$ for $k^* \leq 10$ in the table. This is so because the asymptotic distribution of $\bar{\gamma}$ becomes more skewed as k^* becomes larger, as we see in Figure 1.

The Connor-Korajczyk method to extract factors is based on the homoskedasticity assumption that the aggregate idiosyncratic volatilities, $\frac{1}{n}\varepsilon_t'\varepsilon_t$, where ε_t is the n-vector of idiosyncratic risk, have the same limit for all t within each subperiod. Jones (2001) argues that such an assumption might be violated by more recent data from the US and develops a methodology that relaxes the homoskedasticity assumption. To examine the robustness of the results presented in Table 2, the Jones method is adopted. The same test statistics are calculated as those in Table 2 with the only difference being the way the factors are extracted. The results are reported in Table 3.

Table 3 here

The results from Table 3 show that the conclusion from Table 2 based on the Connor-Korajczyk method is still supported by the results using the Jones method. Generally speaking, the estimated γ and s are even smaller than those in Table 2. One observation from Table 3 is that the statistic $\bar{\gamma}$ is no longer increasing in k^*. This is because, strictly speaking, the Jones method is valid only for the "true" k^*. This lack of monotonicity does not concern us here because one of the $\bar{\gamma}$ is correct (so long as the true k^* is less than 10). As long as no $\bar{\gamma}$ using the Jones method provides evidence against the no-arbitrage implication, the conclusion still holds.

To put the estimated values of the maximum squared Sharpe ratio for each subperiod in perspective, let's look at the sample squared Sharpe ratio of the value-weighted and

equally weighted market portfolio of all the NYSE/AMEX/NASDAQ stocks, as reported in Table 4. Over the eight subperiods from 1965 to 2004, the sample Sharpe ratios vary considerably. For the majority of the subperiods, the squared Sharpe ratio of the value-weighted market portfolio is smaller than that of the equally weighted market portfolio, which, in turn, is close to $\bar{s}$ for $k^* = 1$ in Table 2. In a few subperiods, however, this is reversed.[1]

Table 4 here

III. Simulation Results

Since the test conducted in the last section is based on asymptotic distributions, issues about the difference between finite sample distributions and asymptotic distributions and issues regarding the power of the tests naturally arise. To determine whether we find no violation of the implication of the asymptotic APT because the APT holds well or because the test is not powerful enough, we resort to simulation.

The simulation procedure is explained as follows. The parameters used for the simulation are taken from actual data for their relevance. We take the return data, R, from the subperiod 2000-2004, with $n = 3708$ stocks and $\tau = 60$ months with non-missing values. The variance matrix of the returns is calculated as $(R-\bar{R}1_\tau)(R-\bar{R}1_\tau)'/\tau$ where $\bar{R}$ is the n-vector of mean returns. To generate returns, we take the five eigenvectors of the $\tilde{\Omega} = (R - \bar{R}1_\tau)'(R - \bar{R}1_\tau)/n$ corresponding to the largest five eigenvalues as realizations of three demeaned factors, F. We then regress R on $[1_\tau, F]$ to obtain the the pricing error and the beta matrix, $[a_f, B_F]$. The residuals of the regression are used to calculate its variance matrix, Σ_e. The sample variance matrix of F is denoted as

[1] $\bar{s}$ for $k^* = 1$ is the maximum squared Sharpe ratio, so it should not be exceeded by that of any fixed-weight portfolios. There are two reasons for $\bar{s}$ for $k^* = 1$ to be occasionally smaller than the squared Sharpe ratio of the market portfolios. One is that the stocks used for calculating $\bar{s}$ are those with complete observations in the 60-month subperiods, not all the stocks. The other is that market portfolios do not have fixed-weights as the value of stocks changes month to month.

Σ_F and μ_F is defined as $\mu_F = (B_F'B_F)^{-1}B_F'\bar{R}$. The vector a_F is discarded. Instead, a cross-sectionally normally distributed n-vector, is generated and regressed on B_F to obtain the residual. The pricing error, a, is the residual multiplied by a constant to achieve certain magnitude of mis-pricing. The choice of the constant will be explained below. An $n \times \tau$ return matrix is generated by

$$R = a1_\tau + B_F(\mu_F 1_\tau + \Sigma_F^{1/2}\eta) + \Sigma_e^{1/2}\xi, \tag{8}$$

where η is a $3 \times \tau$ matrix of simulated independent standard normal variables, ξ is an $n \times \tau$ matrix of simulated independent standard normal variables, and η and ξ are independent of each other.

The reason the factor structure is obtained from the sample variance matrix of the actual returns, rather than from the second moment matrix, is to ensure the potential pricing error does not enter the factor construction. It should be obvious that the five-factor structure does not play any important role in generating returns. The variance matrix of the simulated returns is the same as the sample matrix of the actual data from the 2000-2004 subperiod. The only purpose of singling out five factors is to have a B_F matrix to which the pricing error is orthogonal, so that the magnitude of the pricing error can be more meaningful.

In the simulation, we consider three choices of the magnitude of the pricing error. More specifically, we choose the pricing error, a, such that the pricing error per stock, $\sqrt{a'a/n}$, equals 0.02, 0.01 or 0.002. As a very crude estimate, 0.02 is the cross-sectional standard deviation of average monthly excess returns. The choice of $\sqrt{a'a/n} = 0.02$ is interpreted as a large pricing error, the choice of $\sqrt{a'a/n} = 0.01$ can be regarded as a medium-size pricing error, and the choice of $\sqrt{a'a/n} = 0.002$ corresponds to a small pricing error. For each choice of a, one thousand replications are generated. In each replication, i, an $n \times \tau$ matrix of excess returns is generated according to (8), and the statistic $\bar{\gamma}^{(i)}$ is calculated, for a given $k = 1, 2, \cdots, 10$, using the CK method. Correspondingly, the $\bar{s}^{(i)}$ and the left-tail p-values, $p_{0.9}^{(i)}$, $p_{0.5}^{(i)}$ and $p_{0.2}^{(i)}$, are calculated.

The average value of these statistics are then reported in Table 5. Since the returns are generated from homoskedastic distributions, the Jones method is not used here.

Table 5 here

The results in Table 5 show how the magnitude of the test statistics and the left-tail p-values decrease with the magnitude of the pricing error. For the pricing error a with $\sqrt{a'a/n} = 0.02$, the test statistics $\bar{\gamma}$ and $\bar{s}$ appear much greater than the numbers reported in Tables 2 and 3 with actual data. The hypothesis $\gamma \geq 0.9$, however, is rejected for all ks. The left-tail p-values for the hypothesis $\gamma \geq 0.5$ is greater than 1% for $k \geq 7$ and the left-tail p-values for $\gamma \geq 0.2$ are greater than 10% for all $k \geq 3$.

For the pricing error, a, with $\sqrt{a'a/n} = 0.01$, the hypotheses $\gamma \geq 0.9$ and $\gamma \geq 0.5$ are rejected for all ks. The hypothesis $\gamma \geq 0.2$ is not rejected for all ks at the 1% significance level. For the pricing error a with $\sqrt{a'a/n} = 0.002$, the hypotheses $\gamma \geq 0.9$ and $\gamma \geq 0.5$ are rejected, but the hypothesis $\gamma \geq 0.2$ has p-values greater than 1%. Given that the test statistics and the left-tail p-values reported in Tables 2 and 3 are smaller than those in the simulation for the case $\sqrt{a'a/n} = 0.002$, the empirical results in Tables 2 and 3 can be interpreted as the evidence that the actual pricing error is smaller than 0.002 per month on average.

Simulations based on parameters that give a more clear-cut factor structure than that in the US return data are also conducted. The quantitative results are not reported, but the qualitative aspects of the results are described here. In the situation where the number of dominant eigenvalues of the second moment matrix can be more easily determined, the eigenvalues that represent factor variances are large, while the eigenvalues that represent idiosyncratic risk are small. Any pricing error whose squared norm is distinctively greater than those small eigenvalues can be easily detected. Research in international financial markets has found that stock markets in most countries

have more clear-cut factor structures than that in the US.[2] The method of testing the asymptotic APT can be more effective when applied to such markets.

IV. Conclusions

The existing studies on beta-pricing models either test the beta-pricing restriction with pre-specified systematic factors or extract factors from returns without testing the validity of the beta-pricing restriction. Both have shortcomings. Tests with pre-specified factors can be, and in many cases are, inconclusive for the general beta-pricing principle. On the other hand, inferences based on extracted, but contaminated, factors could lead to erroneous conclusions in asset pricing applications. This paper fills the void by developing a test of the beta-pricing restriction using factors extracted from returns.

Based on the asymptotic arbitrage pricing theory advanced by Ross (1976) and Chamberlain and Rothschild (1983), the test proposed in this paper avoids mis-specification of systematic factors by extracting factors from returns using the method proposed by Connor and Korajczyk (1986,1988). The idea is that, if returns are driven by a number of systematic factors and a systematic pricing error, then the number of unbounded eigenvalues in the second-moment matrix of the excess returns will equal the number of unbounded eigenvalues of the variance matrix of the excess returns plus one. In order to be less reliant on identifying the number of factors, an unsettled issue in the literature, the test begins with an arbitrary, but reasonable, number of eigenvectors of the second-moment matrix of returns corresponding to the largest eigenvalues. The smallest eigenvalue of the sample variance matrix of these eigenvectors converges to zero if the asymptotic APT fails. It follows a non-central beta distribution asymptotically, with the non-centrality parameter proportional to the theoretical maximum squared Sharpe ratio of all portfolios of the returns.

The testing method is applied to individual stock returns in the US market during

[2]See, for example, Morck et al. (2000).

eight 60-month subperiods from 1965 to 2004. The empirical evidence supports the implication of the asymptotic APT. The results are robust to the heteroskedastic idiosyncratic risk as the modified test using Jones' (2001) method produces similar results. Simulation experiments show that the test has power to detect large pricing errors, so the results based on actual US return data are indeed evidence of negligible unconditional deviation from the beta pricing principle.

Factors extracted from stock returns are often used to examine asset pricing anomalies associated with firm-specific variables. With the exception of Stambaugh (1983), however, all the literature on APT is cast in the framework of unconditional factor betas. Extending the unconditional APT to a conditional framework and developing tests of the conditional APT remain important tasks for future work.

Appendix

Proof of Proposition 1. Under the assumption of an approximate k-factor structure, $r_t = a + B_f f_t + \varepsilon_t$, where the variance matrix of $B_f f_t$, $B_f \Sigma_f B_f'$, has k unbounded eigenvalues while the variance matrix of ε_t, Σ_ε, has bounded eigenvalues. Let $g_t = S_f^{-1/2} f_t$ and $B_g = B_f S_f^{1/2}$ where $S_f = E f_t f_t'$ is the second-moment matrix of f_t. By construction, $E g_t g_t' = I_k$. Consider the second moment of $a + B_g g_t$, denoted S.

$$S = E[(a + B_g g_t)(a + B_g g_t)'] = aa' + B_g B_g' = (a \ B_g)(a \ B_g)'.$$

The positive eigenvalues of S are the same as those of $Q \equiv (a \ B_g)'(a \ B_g)$. From $a' B_g = a' B_f S_f^{1/2} = 0_k'$, it follows that

$$Q = \begin{pmatrix} a'a & 0_k' \\ 0_k & B_g' B_g \end{pmatrix}.$$

The characteristic function of Q equals

$$|Q - \lambda I_{k+1}| = |a'a - \lambda| \cdot |B_g' B_g - \lambda I_k|.$$

The eigenvalues of Q are therefore $a'a$ and the eigenvalues of $B_g' B_g$. The eigenvalues of $B_g' B_g$ are the same as the positive eigenvalues of $B_g B_g' = B_f S_f B_f' > B_f \Sigma_f B_f'$.[3] Therefore, S has k unbounded eigenvalues if $a'a$ is bounded and S has $k+1$ unbounded eigenvalues if $a'a$ is unbounded. The same is true to the second moment matrix of r, S_r. *Q.E.D.*

Proof of Proposition 2. (i) It is obvious that $\mu_g' \mu_g \geq 0$. Expanding the determinant below in two ways gives

$$\begin{vmatrix} I_{k^*} & \mu_g \\ \mu_g' & 1 \end{vmatrix} = |I_{k^*} - \mu_g \mu_g'| = 1 - \mu_g' \mu_g.$$

Since $I_{k^*} - \mu_g \mu_g'$ is the variance matrix of g, its determinant is non-negative. This establishes that $\mu_g' \mu_g \leq 1$.

[3]For two $n \times n$ symmetric matrices, A_1 and A_2, $A_1 > A_2$ means that $A_1 - A_2$ is positive definite. This implies that the ith largest eigenvalue of A_1 is greater than that of A_2, for $i = 1, \cdots, n$.

(ii) ($\Rightarrow$) When there is a systematic pricing error, $g_t = C(1 \ f_t')'$. Let c' be the first row of C^{-1}. Then $c'g_t = 1$. Since $Eg_tg_t' = I_{k^*}$, it follows that $c'c = 1$. On the other hand, $c = I_{k^*}c = (Eg_tg_t')c = E[g_t(g_t'c)] = Eg_t = \mu_g$. Therefore, $\mu_g'\mu_g = c'c = 1$. ($\Leftarrow$) If $\mu_g'\mu_g = 1$, then obviously $\mu_g \neq 0_{k^*}$. Since $\text{Var}(\mu_g'g_t) = \mu_g'(I_{k^*} - \mu_g\mu_g')\mu_g = 0$, $\mu_g'g_t$ is a constant, and, therefore, the original f_t^* contains a constant, which corresponds to a systematic pricing error.

(iii) Now consider the maximum squared Sharpe ratio in the limit when n goes to infinity. Since $Er_t = a + EBf_t$ and $\text{Var}r_t = \text{Var}(Bf_t) + \text{Var}(\varepsilon_t) > \text{Var}(Bf_t)$, it follows that we only need to consider the linear combinations of the factors. It is the solution to the following problem

$$s = \max_u \frac{[E(u'g_t)]^2}{\text{Var}(u'g_t)} = \frac{u'\mu_g\mu_g'u}{u'(I_{k^*} - \mu_g\mu_g')u}.$$

Here we used the fact that $\text{Var}g_t = Eg_tg_t' - (Eg_t)(Eg_t)' = I_{k^*} - \mu_g\mu_g'$. According to a corollary of the Rayleigh-Ritz principle, the maximum squared Sharpe ratio equals the maximum eigenvalue of $\mu_g\mu_g'$ relative to $I_{k^*} - \mu_g\mu_g'$, which is the maximum solution, s, to

$$\det(\mu_g\mu_g' - s(I_{k^*} - \mu_g\mu_g')) = \det((1+s)\mu_g\mu_g' - sI_{k^*}) = 0.$$

The only positive solution is $s/(1+s) = \mu_g'\mu_g \equiv \gamma$, or

$$s = \frac{\gamma}{1-\gamma}.$$

This completes the proof. *Q.E.D.*

Proof of Proposition 3. (i) From the factor structure, with $\tilde{B} = (a \ B)$ and $\tilde{F} = (1_\tau \ F')'$,

$$\begin{aligned} \Omega &= \tilde{F}'(\frac{1}{n}\tilde{B}'\tilde{B})\tilde{F} + \frac{1}{n}\varepsilon'\varepsilon + \frac{1}{n}\tilde{F}'\tilde{B}'\varepsilon + \frac{1}{n}\varepsilon'\tilde{B}\tilde{F} \\ &\overset{A}{=} H'H + D \end{aligned}$$

where $H = (\text{plim}_{n\to\infty}\frac{1}{n}\tilde{B}'\tilde{B})^{1/2}\tilde{F}$ is a $(k+1)\times\tau$ matrix if $a'a$ is unbounded, or $H = (\text{plim}_{n\to\infty}\frac{1}{n}B'B)^{1/2}F$ is a $k\times\tau$ matrix if $a'a$ is bounded, $D = \text{plim}_{n\to\infty}\frac{1}{n}\varepsilon'\varepsilon = \sigma_\varepsilon I_\tau$ is

a scalar matrix, and $\stackrel{A}{=}$ means asymptotic equivalence. The last two terms in the first line above are zero because ε is uncorrelated with $\tilde{F}$. $H = (h_1, h_2, \cdots, h_\tau)$ is unique up to a pre-multiplier of an orthogonal matrix. Let $\tau^{-1/2}G'$ be the k^* orthonormal eigenvectors of Ω corresponding to the k^* largest eigenvalues of Ω, which is also the eigenvectors of $H'H$. The components of $G = (g_1, \cdots, g_\tau)$ have their sample second-moment matrix equal to I_{k^*} because, by construction, $\tau^{-1}G'G = I_{k^*}$. Without loss of generality, $\tau^{-1/2}G = (HH')^{-1/2}H$. Otherwise, there exists an $\tilde{H} = PH$, where P is orthogonal, such that $\tau^{-1/2}G = (\tilde{H}\tilde{H}')^{-1/2}\tilde{H}$ and we can proceed with $\tilde{H}$. This is inconsequential because H is unique up to an orthogonal multiplier anyway. Note that

$$\tau^{-1}(G - \bar{g}1_\tau')(G - \bar{g}1_\tau')' = (\tau^{-1/2}G)[I_\tau - 1_\tau(1_\tau'1_\tau)^{-1}1_\tau'](\tau^{-1/2}G)',$$

where $I_\tau - 1_\tau(1_\tau'1_\tau)^{-1}1_\tau'$ is a symmetric idempotent matrix with $\tau - 1$ eigenvalues equal to one and one eigenvalue equal to zero. Since $\tau^{-1/2}G'$ comprises orthonormal vectors, it follows from the Poincare separation (interlacing) theorem (see, for example, Horn and Johnson (1985), page 190) that $\tau^{-1}(G - \bar{g}1_\tau')(G - \bar{g}1_\tau')'$ has $k^* - 1$ eigenvalues equal to one and one eigenvalue between zero and one. The smallest eigenvalue equals

$$\min_{u'u=1} \tau^{-1}u'(G - \bar{g}1_\tau')(G - \bar{g}1_\tau')'u = 1 - \max_{u'u=1} u'\bar{g}\bar{g}'u = 1 - (\bar{g}'\bar{g}),$$

which is obtained at $u = \bar{g}/\sqrt{\bar{g}'\bar{g}}$.

(ii) Write

$$\begin{aligned}\bar{\gamma} &\equiv (\bar{g}'\bar{g}) = \tau^{-2}1_\tau'G'G1_\tau = \tau^{-1}1_\tau'H(HH')^{-1}H'1_\tau = \bar{h}'\left(\frac{1}{\tau}\sum_{t=1}^{\tau} h_t h_t'\right)^{-1}\bar{h} \\ &= \bar{h}'\left(\hat{\Sigma}_h + \bar{h}\bar{h}'\right)^{-1}\bar{h} = \frac{\bar{h}'\hat{\Sigma}_h^{-1}\bar{h}}{1 + \bar{h}'\hat{\Sigma}_h^{-1}\bar{h}},\end{aligned}$$

where $\bar{h} = \frac{1}{\tau}\sum_{t=1}^{\tau} h_t$, $\hat{\Sigma}_h = \frac{1}{\tau}\sum_{t=1}^{\tau}(h_t - \bar{h})(h_t - \bar{h})'$ and the last equality is obtained using a matrix inversion formula. It is well known that, under the assumption of normality, $\bar{h}'\hat{\Sigma}_h^{-1}\bar{h}$ is the so-called noncentral Hotelling's T^2 variate, and $\eta \equiv \frac{\tau - k^*}{k^*}\bar{h}'\hat{\Sigma}_h^{-1}\bar{h}$ has a non-central $F(\tau s)$ distribution where $s = \mu_h\Sigma_h^{-1}\mu_h = \mu_{f^*}\Sigma_{f^*}^{-1}\mu_{f^*}$ because h is just a

nonsingular transformation of f^*. See, for example, Muirhead (1982, p.98). The density function of η is

$$p_\eta(y) = \frac{\Gamma(\frac{\tau}{2})}{\Gamma(\frac{\tau-k^*}{2})\Gamma(\frac{k^*}{2})} \frac{y^{k^*/2}\left(\frac{k^*}{\tau-k^*}\right)^{k^*/2}}{\left(1+\frac{k^*}{\tau-k^*}y\right)^{\tau/2}} e^{-\tau s}{}_1F_1\left(\frac{\tau}{2}, \frac{k^*}{2}, \frac{k^* s y}{2(\tau - k^* + k^* y)}\right), \qquad y \geq 0.$$

The change of variable $y = (\tau - k^*)x/[k^*(1-x)]$ with the Jacobian equal to $dy/dx = (\tau - k^*)/[k^*(1-x)^2]$ leads to the density function of $\bar{\gamma}$ as

$$p_{\bar{\gamma}}(x) = \frac{\Gamma(\frac{\tau}{2})}{\Gamma(\frac{\tau-k^*}{2})\Gamma(\frac{k^*}{2})} x^{k^*/2-1}(1-x)^{(\tau-k^*)/2-1} e^{-\tau s/2}{}_1F_1\left(\frac{\tau}{2}, \frac{k^*}{2}, \frac{\tau s x}{2}\right), \qquad 0 \leq x \leq 1.$$

This completes the proof. *Q.E.D.*

References

Bai, J. and Ng, S., 2002, Determining the number of factors in approximate factor models, *Econometrica* 70, 191-221.

Brennan, M., Chordia, T., and Subrahmanyam, A., 1998, Alternative factor specifications, security characteristics, and the cross-section of expected stock return, *Journal of Financial Economics* 49, 345-374.

Chamberlain, G. and Rothschild, M., 1983, Arbitrage, factor structure and mean-variance analysis on large asset markets, *Econometrica* 51, 1281–1304.

Brown, S., 1989, The number of factors in security returns, *Journal of Finance* 44, 1247-1262.

Connor, G. and Korajczyk, R., 1986, Performance measurement with the arbitrage pricing theory: a new framework for analysis, *Journal of Financial Economics* 15, 373-394.

Connor, G. and Korajczyk, R., 1988, Risk and return in an equilibrium APT: application of a new testing methodology, *Journal of Financial Economics* 21, 255-289.

Connor, G. and Korajczyk, R., 1993, A test for the number of factors in an approximate factor model, *Journal of Finance* 48, 1263-1291.

Geweke, J. and Zhou, G., 1996, Measuring the pricing error of the arbitrage pricing theory, *Review of Financial Studies* 9, 557–587.

Gibbons, M., Ross, S. and Shanken, J. 1989, A test of the efficiency of a given portfolio, *Econometrica* 57, 1121-52.

Horn, R.A. and Johnson, C.R., 1985, *Matrix Analysis*, Cambridge University Press.

McCulloch, R. and Rossi, P., 1991, A Bayesian approaches to testing the arbitrage pricing theory, *Journal of Econometrics* 49, 141–168.

Morck, R., Yeung, B. and Yu, W., 2000, The information content of stock markets: why do emerging markets have synchronous stock price movements? *Journal of Financial Economics* 58, 215-260.

Muirhead, R.J., 1982, *Aspects of Multivariate Statistical Theory*, John Wiley & Sons, Inc.

Ross, Stephen, 1976, The arbitrage theory of capital asset pricing, *Journal of Economic Theory* 13, 341–360.

Shanken, J., 1992, The current stage of the arbitrage pricing theory, *Journal of Finance* 47, 1569-1574.

Stambaugh, R., 1983, Arbitrage pricing with information, *Journal of Financial Economics*, 12, 357-370.

Trzcinka, C., 1986, On the number of factors in the arbitrage pricing model, *Journal of Finance* 41, 347-368.

Table 1
Eigenvalue decomposition of the second-moment matrix of individual stocks

This table reports the second moment decomposition of individual stocks in eight 60-month periods. n is the number of stocks. l_j is the jth largest eigenvalue of the second moment matrix of excess returns. $L = \sum_{i=1}^{60} l_i$.

Period	n	L	j	l_j	$\sum_{i=1}^{j} l_i$	l_j/L	$\sum_{i=1}^{j} l_i/L$
1965–1969	1487	21.0684	1	6.077	6.077	0.288	0.288
			2	0.851	6.928	0.040	0.329
			3	0.656	7.584	0.031	0.360
			4	0.585	8.169	0.028	0.388
			5	0.529	8.698	0.025	0.413
			6	0.469	9.167	0.022	0.435
			7	0.455	9.621	0.022	0.457
			8	0.447	10.068	0.021	0.478
			9	0.414	10.483	0.020	0.498
			10	0.394	10.876	0.019	0.516
1970–1974	1824	29.4620	1	10.027	10.027	0.340	0.340
			2	1.613	11.640	0.055	0.395
			3	0.960	12.600	0.033	0.428
			4	0.836	13.436	0.028	0.456
			5	0.628	14.064	0.021	0.477
			6	0.563	14.627	0.019	0.496
			7	0.538	15.165	0.018	0.515
			8	0.514	15.679	0.017	0.532
			9	0.490	16.169	0.017	0.549
			10	0.477	16.646	0.016	0.565
1975–1979	3045	57.8047	1	17.380	17.380	0.301	0.301
			2	2.642	20.022	0.046	0.346
			3	1.761	21.784	0.030	0.377
			4	1.715	23.498	0.030	0.407
			5	1.476	24.974	0.026	0.432
			6	1.362	26.337	0.024	0.456
			7	1.264	27.601	0.022	0.477
			8	1.173	28.774	0.020	0.498
			9	1.100	29.874	0.019	0.517
			10	1.032	30.906	0.018	0.535
1980–1984	3063	55.1144	1	11.400	11.400	0.207	0.207
			2	2.531	13.931	0.046	0.253
			3	1.753	15.684	0.032	0.285
			4	1.566	17.251	0.028	0.313
			5	1.393	18.643	0.025	0.338
			6	1.336	19.980	0.024	0.363
			7	1.266	21.246	0.023	0.385
			8	1.209	22.455	0.022	0.407
			9	1.109	23.564	0.020	0.428
			10	1.059	24.623	0.019	0.447

Table 1 (cont'd)

Period	n	L	j	l_j	$\sum_{i=1}^{j} l_i$	l_j/L	$\sum_{i=1}^{j} l_i/L$
1985–1989	3352	73.6322	1	12.176	12.176	0.165	0.165
			2	2.284	14.460	0.031	0.196
			3	2.061	16.521	0.028	0.224
			4	2.003	18.524	0.027	0.252
			5	1.888	20.412	0.026	0.277
			6	1.836	22.248	0.025	0.302
			7	1.731	23.979	0.024	0.326
			8	1.668	25.647	0.023	0.348
			9	1.602	27.249	0.022	0.370
			10	1.473	28.722	0.020	0.390
1990–1994	3812	115.4322	1	18.010	18.010	0.156	0.156
			2	8.233	26.243	0.071	0.227
			3	3.829	30.072	0.033	0.261
			4	3.460	33.532	0.030	0.290
			5	3.190	36.722	0.028	0.318
			6	2.909	39.631	0.025	0.343
			7	2.807	42.438	0.024	0.368
			8	2.741	45.179	0.024	0.391
			9	2.458	47.638	0.021	0.413
			10	2.437	50.075	0.021	0.434
1995–1999	3858	116.1262	1	14.115	14.115	0.122	0.122
			2	5.995	20.109	0.052	0.173
			3	5.689	25.798	0.049	0.222
			4	5.098	30.896	0.044	0.266
			5	4.386	35.282	0.038	0.304
			6	3.019	38.301	0.026	0.330
			7	2.895	41.195	0.025	0.355
			8	2.788	43.983	0.024	0.379
			9	2.605	46.588	0.022	0.401
			10	2.438	49.026	0.021	0.422
2000–2004	3708	139.2781	1	30.214	30.214	0.217	0.217
			2	9.830	40.044	0.071	0.288
			3	6.695	46.739	0.048	0.336
			4	4.246	50.985	0.030	0.366
			5	3.908	54.893	0.028	0.394
			6	3.229	58.122	0.023	0.417
			7	3.098	61.220	0.022	0.440
			8	2.880	64.101	0.021	0.460
			9	2.789	66.889	0.020	0.480
			10	2.640	69.530	0.019	0.499

Table 2
Test of asymptotic APT with the CK factors

This table reports the results of testing the asymptotic APT using τ observations of k^* factors $\tau^{-1/2}G = \tau^{-1/2}(g_1, \cdots, g_\tau)$ extracted with the Connor-Korajczyk method from the second moment matrix of the excess returns. $\bar{\gamma} = \bar{g}'\bar{g} = 1 - \lambda_{min}$ where $\bar{g}$ is the k^*-vector mean of G, and λ_{min} is the smallest eigenvalue of $\tau^{-1}(G - \bar{g}1'_\tau)(G - \bar{g}1'_\tau)$. $\bar{s} = \bar{\gamma}/(1-\bar{\gamma})$ is the estimate of the asymptotic maximum squared Sharpe ratio, $s = \gamma/(1-\gamma)$. $p_{0.9}$, $p_{0.5}$ and $p_{0.2}$ are the left-tail p-values of $\bar{\gamma}$ under the hypothesis $\gamma = 0.9$, $\gamma = 0.5$, and $\gamma = 0.2$ which correspond to $s = 9$, $s = 1$ and $s = 0.25$ respectively.

Period	k^*	$\bar{\gamma}$	$\bar{s}$	$p_{0.9}$	$p_{0.5}$	$p_{0.2}$
1965–1969	1	0.052	0.055	0.000	0.000	0.020
	2	0.052	0.055	0.000	0.000	0.012
	3	0.079	0.086	0.000	0.000	0.024
	4	0.088	0.097	0.000	0.000	0.022
	5	0.091	0.100	0.000	0.000	0.016
	6	0.092	0.102	0.000	0.000	0.010
	7	0.162	0.193	0.000	0.000	0.071
	8	0.163	0.194	0.000	0.000	0.053
	9	0.166	0.198	0.000	0.000	0.041
	10	0.172	0.208	0.000	0.000	0.034
1970–1974	1	0.026	0.027	0.000	0.000	0.005
	2	0.058	0.062	0.000	0.000	0.016
	3	0.087	0.096	0.000	0.000	0.033
	4	0.090	0.099	0.000	0.000	0.024
	5	0.091	0.101	0.000	0.000	0.016
	6	0.096	0.106	0.000	0.000	0.012
	7	0.097	0.108	0.000	0.000	0.008
	8	0.099	0.110	0.000	0.000	0.005
	9	0.107	0.120	0.000	0.000	0.005
	10	0.129	0.148	0.000	0.000	0.007
1975–1979	1	0.109	0.123	0.000	0.000	0.122
	2	0.129	0.149	0.000	0.000	0.143
	3	0.131	0.151	0.000	0.000	0.114
	4	0.131	0.151	0.000	0.000	0.084
	5	0.138	0.160	0.000	0.000	0.073
	6	0.149	0.175	0.000	0.000	0.071
	7	0.153	0.181	0.000	0.000	0.057
	8	0.153	0.181	0.000	0.000	0.040
	9	0.154	0.182	0.000	0.000	0.028
	10	0.166	0.199	0.000	0.000	0.028
1980–1984	1	0.021	0.022	0.000	0.000	0.003
	2	0.064	0.068	0.000	0.000	0.021
	3	0.068	0.073	0.000	0.000	0.015
	4	0.074	0.080	0.000	0.000	0.012
	5	0.085	0.093	0.000	0.000	0.012
	6	0.090	0.098	0.000	0.000	0.010
	7	0.101	0.113	0.000	0.000	0.010
	8	0.103	0.115	0.000	0.000	0.006
	9	0.105	0.117	0.000	0.000	0.004
	10	0.105	0.117	0.000	0.000	0.002

Table 2 (cont'd)

Period	k^*	$\bar{\gamma}$	$\bar{s}$	$p_{0.9}$	$p_{0.5}$	$p_{0.2}$
1985–1989	1	0.016	0.016	0.000	0.000	0.002
	2	0.019	0.019	0.000	0.000	0.001
	3	0.042	0.044	0.000	0.000	0.004
	4	0.056	0.059	0.000	0.000	0.005
	5	0.084	0.092	0.000	0.000	0.012
	6	0.084	0.092	0.000	0.000	0.007
	7	0.092	0.101	0.000	0.000	0.006
	8	0.095	0.105	0.000	0.000	0.004
	9	0.096	0.107	0.000	0.000	0.003
	10	0.143	0.167	0.000	0.000	0.013
1990–1994	1	0.032	0.033	0.000	0.000	0.007
	2	0.052	0.055	0.000	0.000	0.012
	3	0.070	0.076	0.000	0.000	0.017
	4	0.070	0.076	0.000	0.000	0.010
	5	0.070	0.076	0.000	0.000	0.006
	6	0.089	0.097	0.000	0.000	0.009
	7	0.114	0.128	0.000	0.000	0.016
	8	0.114	0.129	0.000	0.000	0.010
	9	0.114	0.129	0.000	0.000	0.006
	10	0.116	0.131	0.000	0.000	0.004
1995–1999	1	0.096	0.107	0.000	0.000	0.089
	2	0.097	0.107	0.000	0.000	0.066
	3	0.098	0.108	0.000	0.000	0.047
	4	0.099	0.110	0.000	0.000	0.033
	5	0.099	0.110	0.000	0.000	0.022
	6	0.101	0.113	0.000	0.000	0.015
	7	0.102	0.114	0.000	0.000	0.010
	8	0.103	0.114	0.000	0.000	0.006
	9	0.162	0.194	0.000	0.000	0.036
	10	0.172	0.208	0.000	0.000	0.034
2000–2004	1	0.029	0.030	0.000	0.000	0.006
	2	0.033	0.034	0.000	0.000	0.004
	3	0.033	0.034	0.000	0.000	0.002
	4	0.040	0.041	0.000	0.000	0.002
	5	0.127	0.145	0.000	0.000	0.055
	6	0.172	0.208	0.000	0.000	0.120
	7	0.174	0.210	0.000	0.000	0.095
	8	0.175	0.213	0.000	0.000	0.072
	9	0.178	0.217	0.000	0.000	0.057
	10	0.188	0.231	0.000	0.000	0.053

Table 3
Test of asymptotic APT with the Jones factors

This table reports the results of testing the asymptotic APT using τ observations of k^* factors $\tau^{-1/2}G = \tau^{-1/2}(g_1, \cdots, g_\tau)$ extracted with the Jones method from the second moment matrix of the excess returns. $\bar{\gamma} = \bar{g}'\bar{g} = 1 - \lambda_{min}$ where $\bar{g}$ is the k^*-vector mean of G, and λ_{min} is the smallest eigenvalue of $\tau^{-1}(G - \bar{g}1'_\tau)(G - \bar{g}1'_\tau)$. $\bar{s} = \bar{\gamma}/(1-\bar{\gamma})$ is the estimate of the asymptotic maximum squared Sharpe ratio, $s = \gamma/(1-\gamma)$. $p_{0.9}$, $p_{0.5}$ and $p_{0.2}$ are the left-tail p-values of $\bar{\gamma}$ under the hypothesis $\gamma = 0.9$, $\gamma = 0.5$, and $\gamma = 0.2$ which correspond to $s = 9$, $s = 1$ and $s = 0.25$, respectively.

Period	k^*	$\bar{\gamma}$	$\bar{s}$	$p_{0.9}$	$p_{0.5}$	$p_{0.2}$
1965–1969	1	0.048	0.051	0.000	0.000	0.017
	2	0.052	0.055	0.000	0.000	0.012
	3	0.053	0.056	0.000	0.000	0.007
	4	0.071	0.076	0.000	0.000	0.011
	5	0.112	0.126	0.000	0.000	0.035
	6	0.117	0.132	0.000	0.000	0.027
	7	0.120	0.136	0.000	0.000	0.020
	8	0.147	0.172	0.000	0.000	0.033
	9	0.152	0.179	0.000	0.000	0.027
	10	0.175	0.212	0.000	0.000	0.037
1970–1974	1	0.028	0.029	0.000	0.000	0.005
	2	0.048	0.050	0.000	0.000	0.010
	3	0.043	0.045	0.000	0.000	0.004
	4	0.073	0.079	0.000	0.000	0.012
	5	0.093	0.102	0.000	0.000	0.017
	6	0.089	0.098	0.000	0.000	0.009
	7	0.110	0.123	0.000	0.000	0.014
	8	0.106	0.119	0.000	0.000	0.007
	9	0.123	0.140	0.000	0.000	0.009
	10	0.166	0.199	0.000	0.000	0.028
1975–1979	1	0.113	0.128	0.000	0.000	0.134
	2	0.125	0.143	0.000	0.000	0.132
	3	0.159	0.189	0.000	0.000	0.198
	4	0.144	0.168	0.000	0.000	0.114
	5	0.144	0.168	0.000	0.000	0.085
	6	0.147	0.172	0.000	0.000	0.067
	7	0.150	0.177	0.000	0.000	0.052
	8	0.152	0.179	0.000	0.000	0.039
	9	0.182	0.223	0.000	0.000	0.063
	10	0.182	0.223	0.000	0.000	0.045
1980–1984	1	0.018	0.018	0.000	0.000	0.002
	2	0.093	0.103	0.000	0.000	0.058
	3	0.096	0.106	0.000	0.000	0.044
	4	0.110	0.124	0.000	0.000	0.047
	5	0.112	0.126	0.000	0.000	0.035
	6	0.115	0.130	0.000	0.000	0.026
	7	0.124	0.142	0.000	0.000	0.023
	8	0.127	0.146	0.000	0.000	0.017
	9	0.128	0.147	0.000	0.000	0.011
	10	0.148	0.173	0.000	0.000	0.016

Table 3 (cont'd)

Period	k^*	$\bar{\gamma}$	$\bar{s}$	$p_{0.9}$	$p_{0.5}$	$p_{0.2}$
1985–1989	1	0.014	0.014	0.000	0.000	0.002
	2	0.021	0.022	0.000	0.000	0.001
	3	0.082	0.090	0.000	0.000	0.027
	4	0.130	0.149	0.000	0.000	0.082
	5	0.112	0.126	0.000	0.000	0.035
	6	0.114	0.129	0.000	0.000	0.025
	7	0.134	0.154	0.000	0.000	0.032
	8	0.157	0.187	0.000	0.000	0.045
	9	0.187	0.229	0.000	0.000	0.071
	10	0.194	0.240	0.000	0.000	0.062
1990–1994	1	0.045	0.048	0.000	0.000	0.014
	2	0.058	0.061	0.000	0.000	0.016
	3	0.067	0.071	0.000	0.000	0.015
	4	0.075	0.081	0.000	0.000	0.013
	5	0.088	0.097	0.000	0.000	0.014
	6	0.091	0.100	0.000	0.000	0.010
	7	0.121	0.137	0.000	0.000	0.021
	8	0.123	0.141	0.000	0.000	0.015
	9	0.139	0.161	0.000	0.000	0.017
	10	0.152	0.180	0.000	0.000	0.018
1995–1999	1	0.084	0.092	0.000	0.000	0.064
	2	0.107	0.120	0.000	0.000	0.086
	3	0.109	0.123	0.000	0.000	0.065
	4	0.118	0.133	0.000	0.000	0.060
	5	0.132	0.152	0.000	0.000	0.063
	6	0.142	0.166	0.000	0.000	0.059
	7	0.154	0.182	0.000	0.000	0.058
	8	0.158	0.187	0.000	0.000	0.046
	9	0.140	0.163	0.000	0.000	0.018
	10	0.179	0.218	0.000	0.000	0.041
2000–2004	1	0.027	0.028	0.000	0.000	0.005
	2	0.030	0.031	0.000	0.000	0.003
	3	0.036	0.037	0.000	0.000	0.002
	4	0.126	0.145	0.000	0.000	0.074
	5	0.147	0.173	0.000	0.000	0.092
	6	0.166	0.199	0.000	0.000	0.106
	7	0.176	0.214	0.000	0.000	0.099
	8	0.193	0.240	0.000	0.000	0.109
	9	0.182	0.222	0.000	0.000	0.063
	10	0.215	0.274	0.000	0.000	0.100

Table 4
Sharpe ratio of the market portfolios

This table reports the squared Sharpe ratio of market portfolios during the eight subperiods from 1965 to 2004 and the entire sample period. VW is the value-weighted portfolio of NYSE/AMEX/NASDAQ. EW is the equally weighted portfolio of NYSE/AMEX/NASDAQ. Mean and St.dev. are the time-series mean and standard deviation of the market portfolios. Sharpe and Sq-Sharpe are the Sharpe ratio and squared Sharpe ratio, respectively.

Period	Portfolio	Mean	St.dev.	Sharpe	Sq-Sharpe
1965–1969	VW	0.0018	0.0366	0.0506	0.0026
	EW	0.0121	0.0575	0.2114	0.0447
1970–1974	VW	-0.0072	0.0522	-0.1386	0.0192
	EW	-0.0128	0.0685	-0.1863	0.0347
1975–1979	VW	0.0094	0.0443	0.2117	0.0448
	EW	0.0250	0.0634	0.3945	0.1557
1980–1984	VW	0.0035	0.0470	0.0745	0.0055
	EW	0.0058	0.0565	0.1023	0.0105
1985–1989	VW	0.0101	0.0504	0.1996	0.0398
	EW	0.0044	0.0532	0.0820	0.0067
1990–1994	VW	0.0037	0.0361	0.1034	0.0107
	EW	0.0071	0.0450	0.1576	0.0249
1995–1999	VW	0.0167	0.0413	0.4044	0.1636
	EW	0.0117	0.0484	0.2416	0.0584
2000–2004	VW	-0.0020	0.0493	-0.0410	0.0017
	EW	0.0118	0.0676	0.1751	0.0306
1965-2004	VW	0.0045	0.0452	0.0994	0.0099
	EW	0.0081	0.0585	0.1393	0.0194

Table 5
Simulation results

This table reports the results of simulation with 1000 replications. In each replication, excess returns are generated from the parameters calibrated from the excess returns in the 2000-2004 subperiod, plus a normally distributed pricing error, a, with $\sqrt{a'a/n} = 0.02$, 0.01, or 0.002, where $n = 3708$ is the number of stocks. $\bar{\gamma}$ is the test statistics based on CK factors. $\bar{s} = \bar{\gamma}/(1-\bar{\gamma})$ is the estimated asymptotic maximum squared Sharpe ratio. $p_{0.9}$, $p_{0.5}$ and $p_{0.2}$ are the left-tail p-values of $\bar{\gamma}$ under the hypothesis $\gamma = 0.9$, $\gamma = 0.5$, and $\gamma = 0.2$. The numbers reported in the table are the average across 1000 replications.

A. $\sqrt{a'a/n} = 0.02$

k^*	$\bar{\gamma}$	$\bar{s}$	$p_{0.9}$	$p_{0.5}$	$p_{0.2}$
1	0.046	0.051	0.000	0.000	0.045
2	0.071	0.081	0.000	0.000	0.069
3	0.100	0.118	0.000	0.000	0.107
4	0.146	0.183	0.000	0.001	0.190
5	0.187	0.247	0.000	0.003	0.270
6	0.224	0.311	0.000	0.007	0.341
7	0.260	0.380	0.000	0.013	0.407
8	0.297	0.457	0.000	0.022	0.477
9	0.336	0.546	0.000	0.036	0.553
10	0.374	0.649	0.000	0.059	0.622

B. $\sqrt{a'a/n} = 0.01$

k^*	$\bar{\gamma}$	$\bar{s}$	$p_{0.9}$	$p_{0.5}$	$p_{0.2}$
1	0.043	0.048	0.000	0.000	0.043
2	0.066	0.074	0.000	0.000	0.062
3	0.087	0.100	0.000	0.000	0.080
4	0.113	0.134	0.000	0.000	0.109
5	0.136	0.166	0.000	0.000	0.130
6	0.159	0.199	0.000	0.001	0.154
7	0.180	0.231	0.000	0.001	0.175
8	0.200	0.264	0.000	0.002	0.196
9	0.221	0.299	0.000	0.002	0.215
10	0.241	0.335	0.000	0.002	0.236

C. $\sqrt{a'a/n} = 0.002$

k^*	$\bar{\gamma}$	$\bar{s}$	$p_{0.9}$	$p_{0.5}$	$p_{0.2}$
1	0.044	0.049	0.000	0.000	0.043
2	0.062	0.070	0.000	0.000	0.054
3	0.081	0.093	0.000	0.000	0.066
4	0.103	0.120	0.000	0.000	0.084
5	0.123	0.146	0.000	0.000	0.099
6	0.140	0.170	0.000	0.000	0.108
7	0.159	0.197	0.000	0.000	0.123
8	0.177	0.224	0.000	0.001	0.134
9	0.194	0.252	0.000	0.001	0.146
10	0.211	0.280	0.000	0.001	0.157

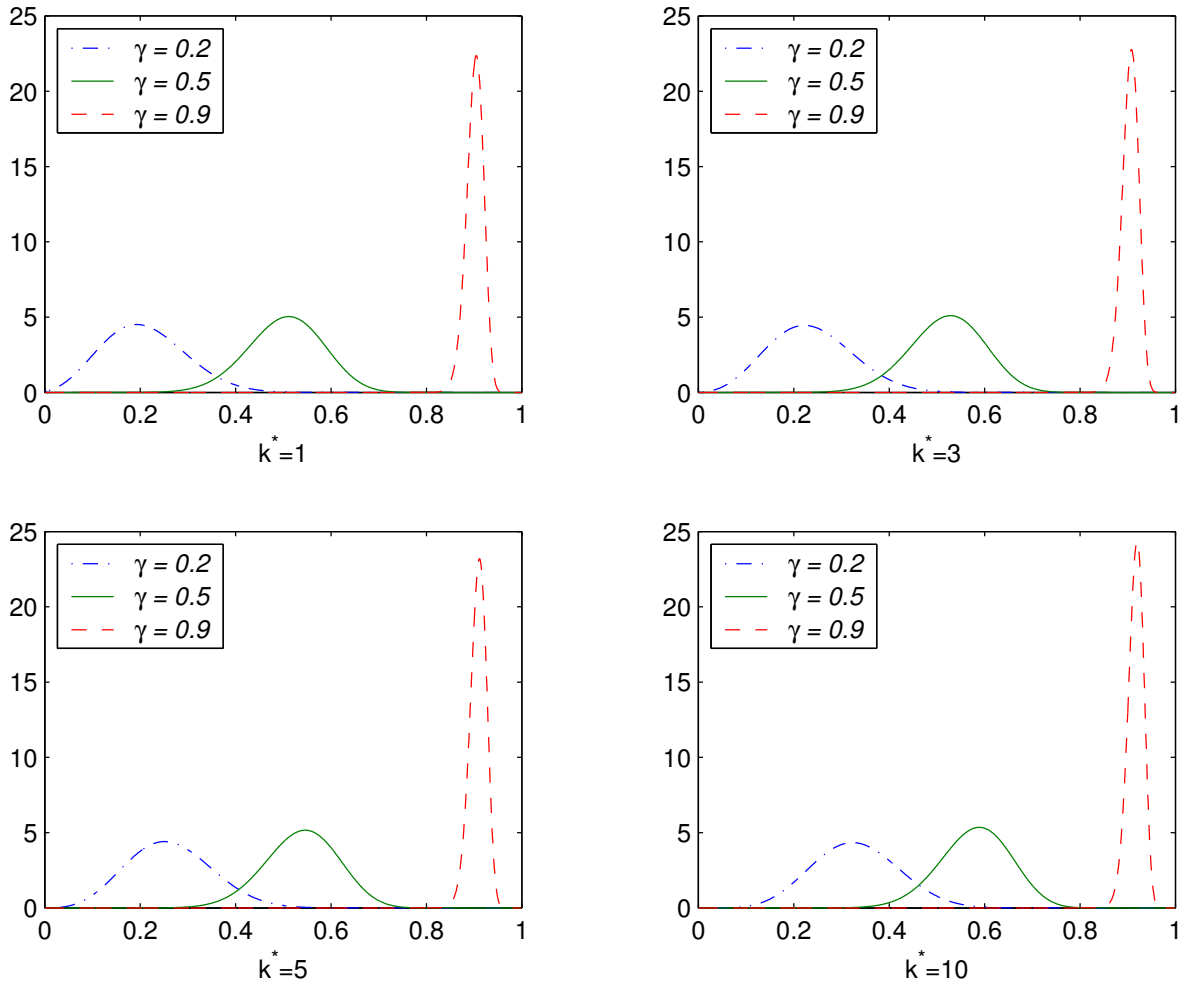

Figure 1. Density functions of noncentral beta distributions.

This figure plots the density functions of noncentral beta distributions in (7) with $\tau = 60$, $k^* = 1, 3, 5, 10$ and $\gamma = 0.2, 0.5, 0.9$.