Deriving the APT when the Number of Factors is Unknown

L. Middleton and S. Satchell

June 1999

L. Middleton is a Research Student at the Judge Institute of Management Studies at the University of Cambridge.

S. Satchell is a member of the Faculty of Economics and Politics and Trinity College, University of Cambridge.

Abstract

This paper examines the use of proxies (or reference variables) for the true factors in the Arbitrage Pricing Theory (APT). It generalises the work of Reisman(1992) and Shanken(1992) and shows that, when there are more reference variables than the true factors, the APT still holds. The possibility of fewer reference variables than the true factors is also considered, but the APT is not shown to hold, in the same sense, for this case. This work builds on an earlier paper by Ingersoll(1984), and our propositions can be thought of as specialisations of his Theorems 1 and 6. Our work does not use the mathematics of Hilbert and Banach spaces (as used by Reisman(1992)) and, thus, is open to a much wider audience.

Deriving the APT when the Number of Factors is Unknown

1) Introduction

The Arbitrage Pricing Theory (APT)*, as proposed by Ross(1976), is an important tool for valuing risky assets. Originally, it assumed that the return from a large number of assets followed a strict factor structure i.e. a linear factor model with uncorrelated disturbances. Ross(1976) does not specify the number or nature of the required factors. Those chosen by researchers range from one to thirteen (and many more for practitioners' models) and there is disagreement as to their identities. It is necessary, therefore, to consider using reference variables i.e. proxies for the true factors.

Shanken(1982) shows that virtually any set of variables can be factors in a well-defined strict factor model. Chamberlain and Rothschild(1983) demonstrate that the Arbitrage Pricing Theory still holds even with an approximate factor model i.e. the disturbances need not be uncorrelated, although the eigenvalues of the residual covariance must be bounded as the number of assets tends to infinity. Ingersoll(1984), Theorem 6, derives a pricing restriction for the case where there are errors in measuring both the expected returns and the matrix of sensitivities to the true set of factors. Our work examines errors in identifying the factors, and, whilst this is mathematically similar to Ingersoll(1984), it is conceptually different. Specifically, we interpret the errors as mis-specifying the nature and number of factors. Thus, we can identify the errors in pricing in terms of characteristics of missing/included factors/proxies.

Reisman(1992), Shanken(1992) and Nawalkha(1997) all investigate the situation where the true factors are unknown but the number of true unknown factors is known. To us, this appears to be implausible for the reasons shown below.

We briefly present a sample from the many authors who have tried to identify the number of true factors from the data. Roll and Ross(1980) estimate that three factors are definitely priced and a fourth may also be present. Dhrymes, Friend and Gultekin(1984) appear to show that the "discovered" factors equate to approximately 10% of the number of the security group analysed, but they do not put it forward as a

proposition, as the exercise was not repeated with many groups. Shanken(1987) finds that there is at least a 99% chance of there being more than ten factors in all cases. Conway and Reinganum(1988) find a cross-validation technique identified one dominant and one minor factor, whilst the likelihood ratio test identified from four to six factors. Connor and Korajczyk(1993) use an asymptotic test for an approximate factor structure and find evidence for one to six pervasive factors. Clare, Priestly and Thomas(1997) show that, by allowing for the existence of an approximate factor structure, five factors, plus a proxy for the market portfolio, are priced in the UK stock market. Chan, Karceski and Lakonishok(1998) find there is nothing to be gained by going beyond the first two or three principal components.

From the above, it can be seen that, to date, research to determine the "correct" number of factors gives a range of one to in excess of ten. This work is further aggravated by the fact that, in constructing the model factors, by using factor analysis and/or principal components, there are possible variations in the process of estimating the factors.

In this paper, we generalise the work of Reisman(1992) and Shanken(1992) by relaxing the assumption that the true number of factors is known. We show that more, or less, reference variables (proxies) can represent the factors. The number of proxies can be the same or more than the "true" factors and the APT remains valid (Propositions 1&2). We also show, in Proposition 3, that less reference variables can also lead to an APT relationship but that the order of magnitude (as a function of the number of stocks) of the squared pricing errors is higher. These orders of magnitude for the two cases of more or less proxies are interpreted in Proposition 4 in terms of the number of pricing errors that exceed a given bound. Our work shows that any number of proxy factors equal to or greater than the true number of factors allows the average pricing error to converge to zero. Apart from the obvious short-comings of factor analysis, this work may also go some way in helping to explain why so many researchers have come up with so many varying numbers of factors. Our proofs avoid the use of the mathematics of Hilbert and Banach spaces and involve only elementary linear algebra.

Section 2 derives general forms of factor structures where proxy factors have replaced the true factors and derives the pricing relations from these factor structures. Section 3 contains some discussion and concluding comments.

2) Factor proxies

This section deduces the general forms of factor structures where proxies have replaced the true factors and, from these, the APT pricing relations are found. The analysis is presented in terms of linear projections, which allow us to present our arguments in terms of matrix algebra.

Consider the following factor structure

$$\mathbf{R} = \mathbf{E} + \mathbf{B}^T\mathbf{T} + \mathbf{e}^T \tag{1}$$

$\mathbf{R}$ is the n-vector of asset returns, $\mathbf{E}$ is the n-vector of expected asset returns, $\mathbf{B}^T$ is the nxm sensitivity matrix, $\mathbf{T}$ is the m-vector of true factors with $E(\mathbf{T}) = 0$ and $cov(\mathbf{T}) = \mathbf{I_m}$ (mxm identity matrix) and $\mathbf{e}^T$ is the n-vector of idiosyncratic risk, $E(\mathbf{e}^T) = 0$ and $cov(\mathbf{e}^T) = \mathbf{D}$, a nxn diagonal matrix.

This gives rise to the pricing relation

$$\mathbf{E} \approx \lambda_0^T\mathbf{i} + \mathbf{B}^T\mathbf{l}^T \tag{2}$$

λ_0^T is a constant, $\mathbf{i}$ is a n-vector of ones, $\mathbf{l}^T$ is a m-vector.

Equation (1) represents the *correct* factor structure, by this we mean that Equation (1) is the true model in the sense that statisticians use the term. Ingersoll(1984), and others take a different stance on what constitutes the true model. He uses the notion of a complete factor representation, (see lemma 4 p1028ff). Needless to say, the true model is a complete factor representation.

Consider a second factor structure

$$\mathbf{R} = \mathbf{E} + \mathbf{B}^P\mathbf{P} + \mathbf{e}^P \tag{3}$$

$\mathbf{R}$ is the n-vector of asset returns, $\mathbf{E}$ is the n-vector of expected asset returns, $\mathbf{B}^P$ is the nxk sensitivity matrix, $\mathbf{P}$ is the k-vector of proxies with $E(\mathbf{P}) = 0$ and $cov(\mathbf{P}) = \mathbf{I_k}$ (kxk identity matrix) and $\mathbf{e}^P$ is the n-vector of mean zero idiosyncratic risk.

If we mistakenly believe that (3) describes the variation in stock prices, can a result similar to (2) be derived? That is, does (4) hold?

$$\mathbf{E} \approx \mathit{l}_0{}^P \mathbf{i} + \mathbf{B}^P \mathbf{l}^P \tag{4}$$

$\lambda_0{}^P$ is a constant, $\mathbf{i}$ is a n-vector of ones, $\mathbf{l}^P$ is a k-vector.

We would like (2) and (4) to be equivalent statements. This requires a relation between $\mathbf{T}$ and $\mathbf{P}$ to be specified. Therefore, let us use the best linear projection (BLP) of $\mathbf{P}$ given $\mathbf{T}$. This can be written as,

$$\mathbf{P} = \mathbf{D}\mathbf{T} + \mathbf{x} \tag{5}$$

$\mathbf{D}$ is a kxm sensitivity matrix and $\mathbf{x}$ is a k-vector of residuals with mean zero.

We wish to substitute the true factors for the proxies, hence, $\mathbf{T}$ needs to be made the subject of Equation (5).

_____

Lemma 1:   *a matrix N exists such that $N\,\mathbf{D} = \mathbf{I}_m$, where $\mathbf{I}_m$ is the (mxm) identity matrix, if and only if rank($\mathbf{D}$) = m.*

Proof – see Appendix – Section 1

_____

This restriction on the existence of $\mathbf{N}$ (Lemma 1) gives two cases that require investigation.

<u>Case (1)</u> – where rank($\mathbf{D}$) = m - ie where k ≥ m

For this case $\mathbf{N}$ exists, and, from now, we will re-label it as $\mathbf{D^{-1}}$. Hence, from (5) we have

$$T = \mathbf{D}^{-1}(P - \mathbf{x}) \qquad (6)$$

Substituting (6) into (1), we have

$$R = E + \mathbf{B^T D}^{-1}(P - \mathbf{x}) + \mathbf{e^T}$$
$$\Rightarrow R = E + \mathbf{B^T D}^{-1}P - \left(\mathbf{B^T D}^{-1}\mathbf{x} - \mathbf{e^T}\right) \qquad (7)$$

<u>Case (2)</u> – where rank($\mathbf{D}$) ≠ m - ie where m > k

For this situation $\mathbf{N}$ does not exist. However, without loss of generality, we can rewrite (5) as,

$$P = \mathbf{D}T + \mathbf{x} = \mathbf{D_1}T_1 + \mathbf{D_2}T_2 + \mathbf{x} \qquad (8)$$

Where,

$$\mathbf{D}T = \mathbf{D_1}T_1 + \mathbf{D_2}T_2$$

and $\mathbf{D_1}$ is non-singular.

Hence, we can now write

$$\mathbf{D_1}^{-1}P = T_1 + \mathbf{D_1}^{-1}\mathbf{D_2}T_2 + \mathbf{D_1}^{-1}\mathbf{x}$$
$$\Rightarrow T_1 = \mathbf{D_1}^{-1}P - \mathbf{D_1}^{-1}\mathbf{D_2}T_2 - \mathbf{D_1}^{-1}\mathbf{x} \qquad (9)$$

Now, without loss of generality, rewrite (1) as

$$R = E + \mathbf{B^T}T + \mathbf{e^T} = E + \mathbf{B_1}^\mathbf{T}T_1 + \mathbf{B_2}^\mathbf{T}T_2 + \mathbf{e^T} \qquad (10)$$

Substituting (9) into (10),

$$\mathbf{R} = \mathbf{E} + \mathbf{B_1}^\mathbf{T}\left(\mathbf{D_1}^{-1}\mathbf{P} - \mathbf{D_1}^{-1}\mathbf{D_2}\mathbf{T_2} - \mathbf{D_1}^{-1}\mathbf{x}\right) + \mathbf{B_2}^\mathbf{T}\mathbf{T_2} + \mathbf{e}^\mathbf{T}$$

$$\Rightarrow \mathbf{R} = \mathbf{E} + \mathbf{B_1}^\mathbf{T}\mathbf{D_1}^{-1}\mathbf{P} + \left(\mathbf{B_2}^\mathbf{T} - \mathbf{B_1}^\mathbf{T}\mathbf{D_1}^{-1}\mathbf{D_2}\right)\mathbf{T_2} + \left(\mathbf{e}^\mathbf{T} - \mathbf{B_1}^\mathbf{T}\mathbf{D_1}^{-1}\mathbf{x}\right) \qquad (11)$$

Note the presence of some of the true factors in the "error" section of this result. This is the general form of equation (7). When the matrix $\mathbf{N}$ exists, $\mathbf{T_2}$ is a matrix of zeroes and $\mathbf{T_1} = \mathbf{T}$ and $\mathbf{B_1}^\mathbf{T} = \mathbf{B}^\mathbf{T}$ and $\mathbf{D_1} = \mathbf{D}$.

_____

A further assumption for the derivation of the APT pricing relation is required. It is that no arbitrage opportunities exist. This is defined in a limiting sense for our proofs and is identical to the asymptotic arbitrage conditions used by Ingersoll(1984) (his Equations (7a) and (7b)). If $\mathbf{w_n}$ is an arbitrage portfolio, then an arbitrage opportunity exists if,

$$\lim{}_{n \to \infty} \mathrm{var}\left(\mathbf{w_n'}\mathbf{R}\right) = 0 \qquad\qquad (12a)$$

$$and \quad \mathbf{w_n'}\mathbf{E} \geq r > 0 \qquad \forall n \qquad\qquad (12b)$$

Proposition 1: Under the assumption of no asymptotic arbitrage, as defined by (12a) and (12b), the factor structure of m true factors can be represented by k proxy factors, correlated with the true factors, where k = m and the APT holds.

Proof:

This is shown by Reisman(1992) Theorems(1) and (2) and by Shanken(1992) when k=m=1.

Proposition 2: Under the assumptions of no asymptotic arbitrage, as defined by (12a) and (12b), and that $E((e^T)^2) \circ s_i^2 < \hat{S}^2$, where S is some bounded constant, the factor structure of m true factors can be represented by k correlated proxy factors where k > m; thus

$$\mathbf{E} \approx \boldsymbol{l}_0\mathbf{i} + \mathbf{B}^*\boldsymbol{l}$$

where $\boldsymbol{B}^* = \boldsymbol{B}^T\boldsymbol{D}^1$. More specifically,

$$\lim_{n \to \infty} \sum_{i=1}^{n} \frac{\left(E_i - \boldsymbol{l}_0 - \sum_{j=1}^{k} B_{ij}^*\boldsymbol{l}_j \right)^2}{n^{2q}} = 0$$

"q > 0; i = 1, 2,..., n; j = 1, 2,..., k

Proof: See Appendix – Section 2

The derived no arbitrage condition is

$$\lim_{n \to \infty} \sum_{i=1}^{n} \frac{\left(E_i - \boldsymbol{l}_0 - \sum_{j=1}^{k} B_{ij}^*\boldsymbol{l}_j \right)^2}{n^{2q}} = 0 \qquad (13)$$

$\forall q > 0$

This result differs from the no arbitrage conditions derived by other researchers (eg Ingersoll(1984) Theorem 1) in that q is usually taken to be ½. This is a minor generalisation. Each term in this infinite sum is non-negative. If q = ½, the average term is zero. Therefore, most of the terms must be negligible. Hence, the linear pricing model, (14), prices most assets correctly, a notion that will be made more precise in Proposition 4. It follows from (13) that

$$\mathbf{E} \approx \boldsymbol{l}_0\mathbf{i} + \mathbf{B}^*\boldsymbol{l} \qquad (14)$$

QED.

We now consider the case where the number of proxies, k, is less than the number of true factors, m.

Proposition 3: Under the assumptions of no asymptotic arbitrage, as defined by (12a) and (12b), and that $E((e^T)^2)$ o $s_i^2 < S^2$, where S is some bounded constant and that the largest value in $\boldsymbol{B_2}^T$ is $(b_2^T)_{max}$ – a bounded constant, the factor structure of m true factors can be represented by k correlated proxy factors where k < m; thus

$$\mathbf{E^T} \approx l_0 \mathbf{i} + \mathbf{B_1^* l}$$

where $\boldsymbol{B_1}^* = \boldsymbol{B_1}^T \boldsymbol{D_l}^{-1}$. More specifically,

$$\lim_{n \to \infty} \sum_{i=1}^{n} \frac{\left(E_i - l_0 - \sum_{j=1}^{k} B_{1ij}^* l_j \right)^2}{n^{2q}} = 0$$

"q > ¹⁄₂, and i = 1, 2,..., n and j = 1, 2,..., k

Proof: See Appendix – Section 3

The derived no arbitrage condition is

$$\lim_{n \to \infty} \sum_{i=1}^{n} \frac{\left(E_i - I_0 - \sum_{j=1}^{k} B_{1ij}^* I_j \right)^2}{n^{2q}} = 0 \qquad (15)$$

$\forall q > ½$

Theorem 1 of Ingersoll(1984) could be applied to the case described in Proposition 3. In which case we would demonstrate that the $\mathbf{W_n}$ in Ingersoll's notation is the cov($\mathbf{f}$) in our notation. We would find a uniform bound of order n (see Equation (a18) in the Appendix), thus we would need $n^{1+\varepsilon}$ for $\varepsilon > 0$ in the denominator of the derived no arbitrage condition to force it to zero.

We now establish how these propositions can give us bounds on the number of large values for v_i.

*Proposition 4: Let **V** be the n-vector with elements v_i. If the number of v_i greater in absolute value than **d** can be written as Cn^q for $C > 0$ and **q** > 0, then given*

$$\lim_{n \to \infty} \sum_{i=1}^{n} \frac{v_i^{\,2}}{n^{2q}} = 0$$

implies that

$$\lim_{n \to \infty} \sum_{i=1}^{n} \frac{v_i^{\,2}}{n^{2q}} \geq \frac{d^2 Cn^q}{n^{2q}} = 0$$

*ie that **q** < 2q*

————

Applying Proposition 4 to Equation (13) shows, for Proposition 2, $\theta < 2q \; \forall q > 0$. Hence, if we say that $q \leq \frac{1}{2}$, the number of mis-pricings (Cn^θ) may not increase as fast as the number of assets. Thus, the mean square pricing error is forced towards zero, as the number of assets becomes large.

However, the result obtained from Equation (15), for Proposition 3, is $\theta < 2q$ $\forall q > \frac{1}{2}$. Therefore, the increase in the number of mis-pricings may be the same as the increase in assets. Thus, the mean square pricing error is not forced to be zero as the number of assets becomes large.

3) - Concluding Comments

Proposition 1, that an equal number of proxy factors can represent the set of true factors was proven in the work of Reisman(1992).

Proposition 2, that a greater number of proxy factors can represent the set of true factors, has resulted in generalising the work of Reisman(1992) and Shanken(1992) without the mathematical complexity of the former.

Proposition 3, that a fewer set of proxies leads to an APT type relationship seems unexpected. However, it is a consequence of Proposition 4 that the APT relationship of Proposition 3 cannot be shown to have a relatively small number of large deviations in the same sense that an APT relationship derived from Proposition 2 can.

Ingersoll(1984), in his theorem 1, examines the case where the covariance matrix of the errors is not assumed to be diagonal. However, he does not address the question as to whether the APT prices have zero average error for this case. In his theorem 6, he shows that the problem where there are errors in measuring both the risk free rate of return and the sensitivity matrix can be reduced to the same problem as in his theorem 1. Our work shows that the problem of errors in identifying the factors can be reduced to the problem in theorem 1. We, however, are able to prove that, when the number of proxy factors is greater or equal to the number of true factors, the APT does price with zero average error. We are unable to prove a similar result for the case where there are fewer proxies than the true factors, but we can give some measure of the average error in this case, see Proposition (4).

The fact that more proxies could represent a fewer number of true factors was expected. The practical implications of this work are that the *correct* APT pricing relation would be best represented by an equal or greater number of correlated proxy factors. Hence, if there is factor uncertainty in the model, and we would like the APT to hold, the principal of parsimony seems inappropriate. Thus, the empirical implication of our paper for model builders is to be generous in the number of factors when building linear factor models.

Appendix

1) <u>Proof of lemma 1</u>

Suppose that the columns of **D** are given by column vectors $v_1, \ldots, v_m$. These vectors are linearly independent so we can extend the set to a basis for k-dimensional space:

$$v_1, \ldots, v_m, v_{m+1}, \ldots, v_k$$

Now the matrix **D′**, resulting from putting all these column vectors side by side, has two properties:

    (i)   it is invertible (because the columns are linearly independent)

    (ii)   the first m columns are the original matrix **D**.

By the first property, we can find a kxk matrix **N′** such that, $\mathbf{N'D'} = \mathbf{I_k}$. And, taking the first m rows of **N′** will give a mxk matrix **N** such that $\mathbf{N\,D} = \mathbf{I_k}$

QED

_____

2) Proof of Proposition 2

Equation (7) states

$$\mathbf{R} = \mathbf{E} + \mathbf{B}^*\mathbf{P} - \left(\mathbf{B}^* - \mathbf{e}^{\mathbf{T}}\right) \tag{7}$$

Where

$$\mathbf{B}^* = \mathbf{B}^{\mathbf{T}}\mathbf{D}^{-1}$$

It follows from the definitions of the various quantities that

$$E\left(\mathbf{B}^*\mathbf{x} - \mathbf{e}^{\mathbf{T}}\right) = 0 , \qquad\qquad E\left(\mathbf{e}^{\mathbf{T}}\mathbf{e}^{\mathbf{T'}}\right) = \mathbf{D}$$

$$E(\mathbf{PP'}) = \mathbf{I_k} , \qquad\qquad E(\mathbf{P}) = 0$$

Regress $\mathbf{E}$ with $\mathbf{B}^*$ and a vector of ones. Call the regression coefficients $\mathbf{l}$ and λ_0 respectively. This should, in general, be possible, as there should be no collinearity. Thus we have

$$\mathbf{E} = l_0\mathbf{i} + \mathbf{B}^*\mathbf{l} + \mathbf{v} \qquad (a1)$$

$\mathbf{v}$ are the residuals of the regression and have the conventional properties of BLP residuals,

$$\mathbf{i}'\mathbf{v} = 0, \qquad (a2)$$

$$\mathbf{B}^{*'}\mathbf{v} = 0 \qquad (a3)$$

Consider the arbitrage portfolio $\mathbf{w}$, with elements

$$w_i = v_i / \left(\|v_n\| n^q \right) \qquad \forall q > 0 \qquad (a4)$$

where

$$\|v_n\|^2 = \sum_{i=1}^{n} v_i^2$$

The profit on this portfolio is

$$\mathbf{w}'\mathbf{R} = \left(\sqrt{n}\|v_n\| \right)^{-1} \mathbf{v}'\mathbf{R} = \left(n^q \|v_n\| \right)^{-1} \mathbf{v}'\left(\mathbf{E} + \mathbf{B}^*\mathbf{P} - \left(\mathbf{B}^*\mathbf{x} - \mathbf{e}^{\mathbf{T}} \right) \right)$$

$$= \left(n^q \|v_n\| \right)^{-1} \mathbf{v}'\left(\mathbf{E} + \mathbf{e}^{\mathbf{T}} \right) \qquad (a5)$$

from (a3)

The expected profit is

$$\mathbf{w'E} = \left(\sqrt{n}\|v_n\|\right)^{-1}\mathbf{v'E} = \left(n^q\|v_n\|\right)^{-1}\mathbf{v'}\left(l_0\mathbf{i} + \mathbf{B}^*\mathbf{1} + \mathbf{v}\right)$$

$$= \frac{\|v_n\|}{n^q} \qquad (a6)$$

from (a2) and (a3)

The variance of the profit is

$$\mathrm{var}\left(\mathbf{w'R}\right) = E\left(\left(\mathbf{w'R} - E(\mathbf{w'R})\right)\left(\mathbf{w'R} - E(\mathbf{w'R})\right)'\right)$$

$$= E\left(\left(\mathbf{w'e^T}\right)\left(\mathbf{w'e^T}\right)'\right)$$

$$= \mathbf{w'Dw} \le \left(n^q\|v_n\|\right)^{-2}S\mathbf{v'v} = \frac{S^2}{n^{2q}} \qquad (a7)$$

If the proposition is false, the expected profit (a6) remains nonzero while the variance would tend to zero as n increases. This would represent an opportunity of the kind prohibited by asymptotic arbitrage. Hence, $\|v_n\|^2/n^{2q}$ must vanish.

The derived no arbitrage condition is

$$\lim_{n\to\infty}\sum_{i=1}^{n}\frac{\left(E_i - l_0 i_i - \sum_{j=1}^{k}B_{ij}^*l_j\right)^2}{n^{2q}} = 0 \qquad (a8)$$

QED

_____

3) Proof of Proposition 3

Equation (11) states

$$\mathbf{R} = \mathbf{E} + \mathbf{B_1^T D_1}^{-1}\mathbf{P} + \left(\mathbf{B_2^T} - \mathbf{B_1^T D_1}^{-1}\mathbf{D_2}\right)\mathbf{r_2} + \left(\mathbf{e^T} - \mathbf{B_1^T D_1}^{-1}\mathbf{x}\right) \qquad (11)$$

Rewrite (11) as

$$\mathbf{R} = \mathbf{E} + \mathbf{B}_1^{*}\mathbf{P} + \mathbf{f} \qquad (a9)$$

where

$$\mathbf{f} = \mathbf{B_2}^{\mathbf{T}}\mathbf{T_2} + \mathbf{e}^{\mathbf{T}} + \mathbf{B}_1^{*}\mathbf{G}$$

$$\mathbf{G} = \left(-\mathbf{D_2}\underline{\mathbf{T_2}} - \mathbf{x}\right)$$

It follows from the definitions of the various quantities that

$$E(\mathbf{f}) = 0 , \qquad\qquad E\!\left(\mathbf{e}^{\mathbf{T}}\mathbf{e}^{\mathbf{T}'}\right) = \mathbf{D}$$

$$E(\mathbf{P}) = 0 , \qquad\qquad E(\mathbf{PP'}) = \mathbf{I_k}$$

The covariance matrix of the errors is given by

$$\operatorname{cov}(\mathbf{f}) = E(\mathbf{ff'})$$
$$\Rightarrow \operatorname{cov}(\mathbf{f}) = \mathbf{B_2^{T}B_2^{T'}} + \mathbf{D} + \mathbf{B_2^{T}}\operatorname{cov}(\mathbf{T_2},\mathbf{G'})\mathbf{B_1^{*'}}$$
$$+ \operatorname{cov}\!\left(\mathbf{e}^{\mathbf{T}},\mathbf{G'}\right)\mathbf{B_1^{*'}} + \mathbf{B_1^{*}}\operatorname{cov}\!\left(\mathbf{G},\mathbf{T_2}'\right)\mathbf{B_2^{T'}} + \mathbf{B_1^{*}}\!\left(\mathbf{G},\mathbf{e}^{\mathbf{T}'}\right) + \mathbf{B_1^{*}}(\mathbf{GG'})\mathbf{B_1^{*'}} \quad (a10)$$

Regress $\mathbf{E}$ with $\mathbf{B_1}^{*}$ and a vector of ones. Call the regression coefficients $\mathbf{l}$ and λ_0, thus we have

$$\mathbf{E} = l_0\mathbf{i} + \mathbf{B_1^{*}l} + \mathbf{v} \qquad (a11)$$

Where $\mathbf{v}$ is the vector of regression residuals, with the usual properties,

$$\mathbf{i'v} = 0, \qquad (a12)$$
$$\mathbf{B_1^{*'}v} = 0 \qquad (a13)$$

Consider a portfolio $\mathbf{w}$ where

$$w_i = \frac{v_i}{\|v_n\|n^q} \qquad\qquad \forall q > \tfrac{1}{2} \qquad (a14)$$

The profit on this portfolio is

$$\mathbf{w}'\mathbf{R} = \mathbf{w}'\mathbf{E} + \mathbf{w}'\mathbf{B}_1^*\mathbf{P} + \mathbf{w}'\mathbf{f}$$

$$\Rightarrow \mathbf{w}'\mathbf{R} = \mathbf{w}'\mathbf{E} + \mathbf{w}'\mathbf{f} \qquad (a15)$$

The expected profit is

$$\mathbf{w}'\mathbf{E} = \frac{\mathbf{v}'}{\|v_n\| n^q} \left(\mathbf{B}_1^*\mathbf{1} + I_0\mathbf{i} + \mathbf{v}\right)$$

$$\Rightarrow \mathbf{w}'\mathbf{E} = \frac{\|v_n\|}{n^q} \qquad (a16)$$

The variance on this profit is

$$\mathrm{var}(\mathbf{w}'\mathbf{R}) = E\left((\mathbf{w}'\mathbf{R} - E(\mathbf{w}'\mathbf{R}))(\mathbf{w}'\mathbf{R} - E(\mathbf{w}'\mathbf{R}))'\right)$$

$$= E\left((\mathbf{w}'\mathbf{f})(\mathbf{w}'\mathbf{f})'\right)$$

$$= \mathbf{w}'\mathrm{cov}(\mathbf{f})\mathbf{w} \qquad (a17)$$

$$\Rightarrow \mathrm{var}(\mathbf{w}'\mathbf{R}) = \frac{\mathbf{v}'\left(\mathbf{B}_2^{\mathbf{T}}\mathbf{B}_2^{\mathbf{T}'} + \mathbf{D}\right)\mathbf{v}}{\|v_n\|^2 n^{2q}} \qquad (a18)$$

From (a12) and (a13)

Therefore we need to find a bound for the numerator.

_____

Lemma 2:

If the n×n matrix A has n distinct eigenvalues and is positive semidefinite, then for some n-vector, v,

$$\mathbf{v}'\mathbf{A}\mathbf{v} \le \mathbf{v}'\mathbf{v}\, I_A^{\max} \le \mathbf{v}'\mathbf{v}\, tr(\mathbf{A}) \le n\mathbf{v}'\mathbf{v}\, a^{\max} \qquad (a19)$$

Where $l_A{}^{max}$ is the largest eigenvalue of A, tr(A) is the trace of A and d^{max} (< ¥) is the largest value in A.

———

Inserting (a19) into (a18) gives

$$\mathrm{var}(\mathbf{w'R}) \le \frac{\mathbf{v'v}n(m-k)\left(b_2^T\right)_{max}^2 + \mathbf{v'Dv}}{\|v_n\|^2 n^{2q}} \qquad (a20)$$

Where $(b_2{}^T)_{max}$ is the largest value in $\mathbf{B_2}^\mathbf{T}$

$$\Rightarrow \mathrm{var}(\mathbf{w'R}) \le \frac{(m-k)\left(b_2^T\right)_{max}^2}{n^{2q-1}} + \frac{S^2}{n^{2q}} \qquad (a21)$$

Assuming that $E((\mathbf{e}^\mathbf{T})^2) \equiv s_i^2 < S^2$.

If the proposition is false, the expected profit remains nonzero (see (a16)) while the variance would tend to zero as n increases. This would represent an opportunity of the kind prohibited by asymptotic arbitrage. Hence, $\|v_n\|^2/(n^{2q})$ must vanish.

The derived no arbitrage condition is

$$\lim_{n\to\infty} \sum_{i=1}^{n} \frac{\left(E_i - l_0 - \sum_{j=1}^{k} B_{1ij}^* l_j\right)^2}{n^{2q}} = 0 \qquad (a22)$$

QED

References

Gary Chamberlain and Michael Rothschild, "Arbitrage, Factor Structure, and Mean-Variance Analysis on Large Asset Markets," Econometrica 51, (Sept 1983), 1281 – 1304

Louis K.C. Chan, Jason Karceski, and Josef Lakonishok, "The Risk and Return from Factors," Journal of Financial and Quantitative Analysis 33, (June 1998), 159-188

Andrew Clare, Richard Priestly and Stephen Thomas, "The Robustness of the APT to Alternative Estimators," Journal of Business Finance and Accounting, (1997), 645 - 655

Gregory Connor and Robert A. Korajczyk, "A Test for the Number of Factors in an Approximate Factor Model," Journal of Finance 48, (September 1993), 1263 - 1291

Gregory Connor and Robert A. Korajczyk, "The Arbitrage Pricing Theory and Multifactor Models of Asset Returns," Chapter 4 of Finance, Handbooks in Operations Research and Management Science, Volume 9, edited by R. Jarrow, V. Maksimovic, and W. Ziemba. Amsterdam: North Holland, (1995)

Delores A. Conway and Marc R. Reinganum, "Stable Factors in Security Returns: Identification Through Cross Validation" Journal of Business & Economic Statistics 6, (January 1988), 1 - 15

Phoebus Dhrymes, Irwin Friend and N. Bulent Gultekin, "A Critical Reexamination of the Empirical Evidence on the Arbitrage Pricing Theory," Journal of Finance, (1984), 323 – 346

Jonathan Ingersoll, "Some Results in the Theory of Arbitrage Pricing," Journal of Finance 39, (1984), 1021 – 1039

J. Ingersoll, "Theory of Financial Decision Making," Rowman & Littlefield, (1987)

Sanjay Nawalkha, "A Multibeta Representation Theorem for Linear Asset Pricing Theories," Journal of Financial Economics 46, (1997), 357 - 381

Haim Reisman, "Reference Variables, Factor Structure, and the Approximate Multibeta Representation," Journal of Finance 47, (September 1992), 1303-1314

Richard Roll and Stephen Ross, "An Empirical Investigation of the Arbitrage Pricing Theory," Journal of Finance 35, (1980), 1073 – 110

Stephen Ross, "The Arbitrage Theory of Capital Asset Pricing," Journal of Economic Theory 13, (1976), 341 – 360

Jay Shanken, "The Arbitrage Pricing Theory : Is it Testable?," Journal of Finance, (1982), 1129 - 1140

Jay Shanken, "Nonsynchronous Data and the Covariance-Factor Structure of Returns," Journal of Finance 42, (June 1987), 221-231

Jay Shanken, "The Current State of the Arbitrage Pricing Theory," Journal of Finance 47, (September 1992), 1569 - 1574