TAPMI

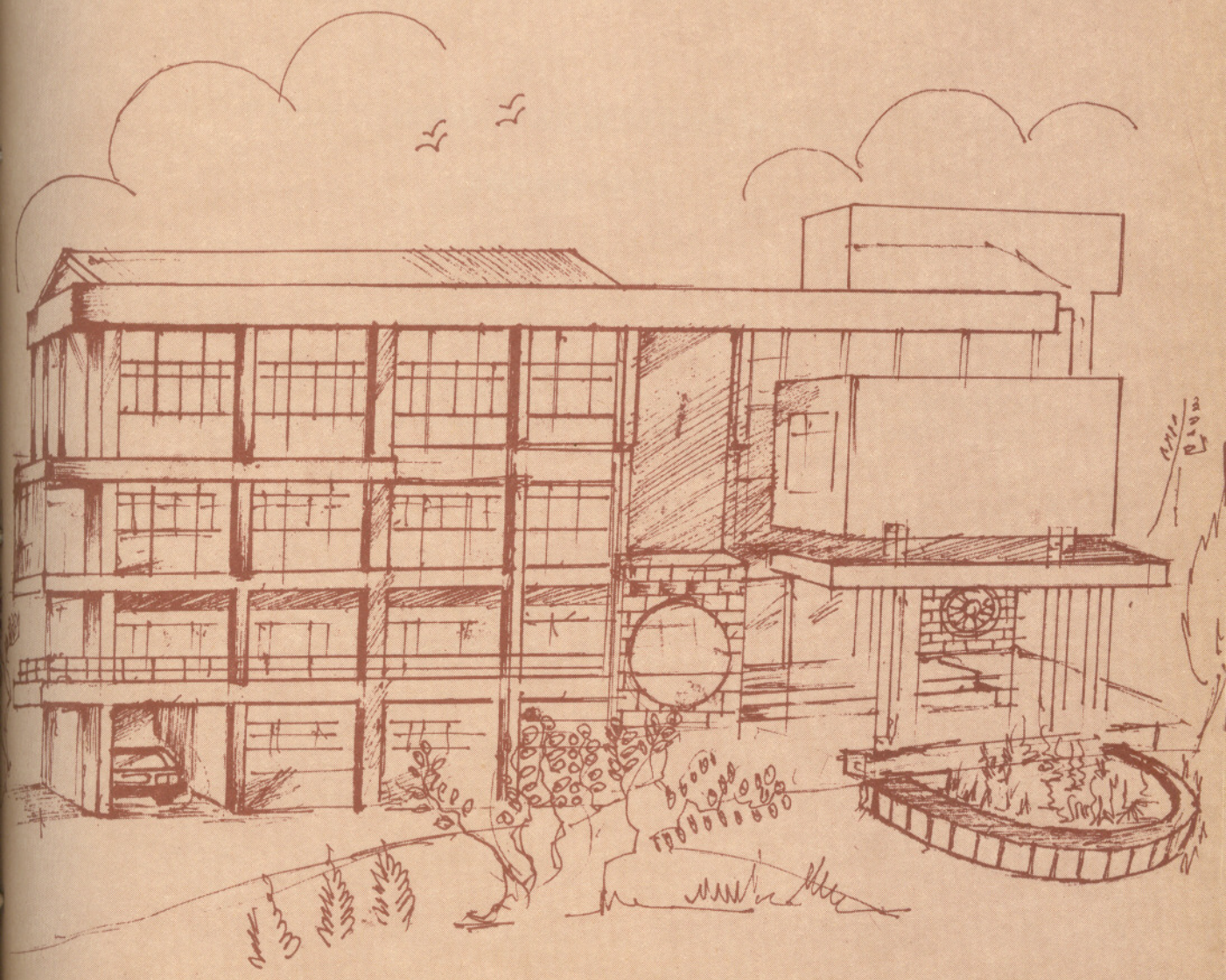
Working Paper Series

No. 12

DOES APT WORK IN INDIA ?

By

PITABAS MOHANTY
Assistant Professor



DOES APT WORK IN INDIA ?

By

PITABAS MOHANTY
Assistant Professor

TAPMI WORKING PAPER SERIES NO. 98/02

February 1998

The Objective of TAPMI working paper series is to help faculty members of TAPMI to test out their research ideas/findings at the pre-publication stage.

T.A.PAI MANAGEMENT INSTITUTE
MANIPAL - 576119

Does APT work in India?[1]

Abstract

No. It does not appear to.

Introduction:

Return generating models play a very valuable role in the theory of Finance. They are used in the estimation of the cost of equity, in portfolio management (for example in creating a portfolio with a particular factor load), in portfolio performance measurement, in testing the efficiency of the capital market, etc. The usage of return generating models (among the practitioners) is however, limited to one of estimation of cost of equity. Financial Institutions and consultancy firms use (mostly) CAPM as the return generating model. In fact, some of the practitioners confided to me that CAPM is used simply because it is very simple to use and not because it is such a great model. Some of them told me that Arbitrage Pricing Theory (APT) is relatively new[2](!) and hence is not used. However, Richard Roll's (1977,1994) criticism of CAPM is well known and hence it is advisable not to use CAPM as the mistakes we make are not systematic and hence we cannot know whether we are over- or under-estimating the expected return of a capital asset.

This working paper briefly discusses APT, reviews the literature on APT, and finally tests APT in the Indian context.

The Arbitrage Pricing Theory

APT is a different approach to determining equilibrium returns of assets. The APT description of equilibrium is more general than that provided by a CAPM-type model. One just requires the homogeneous expectation assumption (among the assumptions required in CAPM). The central core of the APT as outlined by Ross is that the return generating process for a population of N assets is a linear function of only a few, say k ($k \leq N$) systematic factors or components and as a consequence many portfolios are close substitutes. The return generating process can be written in equation form as

$$R(i) = E(i) + b(i,1)*f(1) + \ldots + b(i,k)*f(k) + e(i) \ldots\ldots (2).$$

[1] The author wishes to thank Prof. J R Verma of IIM Ahmedabad for his valuable comments on the paper.

[2] APT was developed as long back as in 1976.

Here

R(i) is the observed return of stock i,

E(i) is the expected return of stock i,

b(i,j) is the factor loading of stock i with respect to factor j,

f(j) is the jth factor score.

We assume that the expected values of the factor scores as well as the idiosyncratic error terms are zero.

So if there are enough securities in the market so that a portfolio can be formed with the following portfolio weights :

x1+x2+...+xN = 0. (This of course, is possible if we assume that short sales is possible. But latter literature has shown that this assumption is not necessary)

x1*b(1,i) + x2*b(2,i) + ... +xN*b(N,i) = 0, i= 1,...,k.

x1*e(1) + ... +xN*e(N) = 0,

then E(i) can be written as a linear combination of the b(i,j)'s and a constant vector. So

E(i) = a(0) + b(i,1)*a(1) + ... + b(i,k)*a(k) ..... (3).

LITERATURE REVIEW

In this section the literature regarding APT (both the theoretical basis as well as the empirical tests) has been discussed briefly. The findings regarding APT are mixed. While some have found out evidence supporting APT (Roll and Ross (1980) and Chen (1985)), others have found that APT is either misspecified (Reinganum (1984) or that APT tests are not robust (Dhrymes, Friend, and Gultekin (1984)).

Chen and Ingersol (1983) found that a sufficient condition for the pricing relation to hold for all assets in the economy when return generating process is given by Eqn(2) is that there exists a portfolio whose return has no unsystematic risk, and that some expected utility maximizer with a continuously differentiable, increasing and strictly concave utility function U and initial wealth W considers that portfolio to be his optimal portfolio. Here we don't require the assumption that there are infinitely many assets in the economy.

Ross (1978) has proved that if returns are described by the linear factor model given by eqn(2), and all investors hold optimal portfolios which are combinations of k+1 residual

risk- free portfolios, then the pricing relation given by eqn(3) must hold. This result, of course, depends on portfolio separation.

In such a setting, the risk-less return and each of the k factors can be expressed as a linear combination of k+1 other factors, say R(1) through R(k+1). Any other asset's return, since it is a linear combination of the factors, must be a linear combination of the first k+1 assets' returns. And thus portfolios of the first k+1 assets are perfect substitutes for all other assets in the market. Since perfect substitutes must be priced equally, there must be restrictions on the individual returns generated by the model.

The first step is to estimate simultaneously the factors (f(j)'s) and firm attributes (b(i,j)'s) from equation (2). A complete specification of Eqn.(2) would call for all factors (f(j)'s) and attributes (b(i,j)'s) to be defined, so that the covariance between any residual return is zero. While it may not always be possible to produce this exact result, factor analysis can be used to produce results that approximate this result. Factor analysis determines a specific set of f(i)'s and b(i,j)'s such that the covariance of residual returns (returns after the influence of these factors have been removed) is as small as possible.

Roll and Ross [3](1980) were the first to have attempted to test the APT. From eqn.(2) they decomposed the population variance, V, into

$V = B L B' + D$ ....(4),

where $B = [b(i,j)]$, is the matrix of factor loadings, L is the matrix of factor covariances, and D is the diagonal matrix of own asset variances.

In an economy with many assets this is basically a requirement that one can approximately duplicate the factors with linear combination of assets, or the requirement that B B' must have k very large eigen values with the remaining eigen values zero.

If G is any orthogonal matrix such that $G G' = I$, then

$V = (B G) (G' L G) (B G)' + D.$

Thus if B is to be estimated from V, then all transforms B G will be equivalent. Thus for example, we can scale up factor j's loadings and scale down j by the same constant

[3] Here this methodology has been discussed at length because this working paper has replicated Roll and Ross methodology without making any change.

g, leaving distribution. of returns unaltered. We can eliminate ambiguity by restricting the factors to be orthonormal so that they are Independent ~~pendent~~ and have unit variances. *The above constraint alters the form of null hypothesis of APT, but will not affect the statistical rejection region.*

Once E(i) and B have been estimated, we can then move to estimate the a(i,j)'s. The general procedure is to examine cross-sectional regressions of the form

E(i) = E(0) + a(i,1)*b(i,1)+ ... + a(i,k)*b(i,k) + error terms,

where E(0), and a(i,j)'s are to be estimated.

Estimating the Factor Model

- For a group of individual assets, a sample product moment covariance matrix computed from a time-series of returns.
- ML factor analysis is performed on the covariance matrix. This estimates the number of factors and the matrix of loadings.
- The individual asset factor loading estimates from the previous step are used to explain the cross-sectional variation of individual expected returns. The procedure here is similar to a cross-sectional generalised least squared regression.
- Estimates from the cross-sectional model are used to measure the size and the statistical significance of risk-premia associated with the estimated factors.
- Roll and Ross divided the entire sample into 42 groups of 30 companies each. The above procedure was repeated for all the 42 groups.

The MLE method also estimates the number of factors. This is accomplished by specification of arbitrary number of factors, say j, then solving for maximum likelihood conditional on a covariance matrix generated by exactly j factors. A second value of the likelihood function will also be found which is conditional on the observed sample covariance matrix without any restriction as to the no. of factors. Then a likelihood ratio (first likelihood divided by the second), is computed. Under the null hypothesis of exactly j factors, twice the natural log of the likelihood ratio is distributed asymptotically as a chi-square with $0.5*[(N-k)^2 - (N-k)]$ degrees of freedom. Thus if the computed chi-square statistic is large (small) then more (fewer) than j factors are required to explain the structure of the return generating process. One can stop when the

probability that the next factor explains a statistically significant portion of the covariance matrix drops below some level, for example, 50%.

Roll and Ross allowed the possibility of spurious factors at stage 2 because the same number of true(priced) factors should be present in every group and the first group might have been unrepresentative. Fewer than the true number of common factors could have been estimated from group 1 because of sampling variation. The third stage protects against too many factors estimated at stage 2, but it does not protect against too few.

Combining Eqn(2) and Eqn(3), we obtain

$r(t) = R(t) - a(0) = B\,L + (B\,F + E)$

$= B\,L + X(t)$, ... (5).

where expected value $X(t) = 0$.

The factor loadings are chosen such that

$V = B\,B' + D$ is the estimated covariance matrix of $B\,F + E$, the disturbance term in Eqn(5). Roll and Ross estimated the risk premia using the formula

$L_t = (B'\,V^{-1}\,B)^{-1}\,B'\,V^{-1}\,r(t)$ ..... (6).

$= T\,r(t)$.

The covariance matrix of the above estimates L_t from eqn(6) can be obtained by

$1/t\,B'\,V^{-1}\,B$ ..... (7).

Here t is the number of time-series data points we have with us.

The above is very convenient to use since it is constrained to be diagonal. So the estimated risk premia are mutually independent and admit simple t-tests of significance. In eqn(6) a value of a(0) has to be assumed.

Roll and Ross also tested for the equivalence of the factor structure across groups. They found no evidence to support the hypothesis that the intercept terms (factor structure) were different across groups.

Chow, Elton, and Gruber (1984) repeated the Roll and Ross methodology for a latter time period to test the robustness of the Roll and Ross methodology. They simulated returns from a zero-beta CAPM. When Roll and Ross methodology was applied to these data, the number of factors that was found to be significant ~~cant~~ was consistent with the zero-beta form of the CAPM. The fact that more factors were found to be

significant when applied to real life data leads support to Roll and Ross's argument that additional factors beyond those embodied by the zero-beta form of the capital assets pricing model determining equilibrium prices.

Reinganum (1984) tested the APT against the alternative hypothesis that the size effect vanishes after the APT risk adjustment is done. His test of APT is broken down into two stages. In year T-1, factor loadings are estimated for all securities. Securities with similar factor loadings are grouped into control portfolios. For example, with a three-factor model, sixty-four (4*4*4) control portfolios were formed. Membership in one of the sixty- four portfolios is determined based on the magnitude of the factor loadings.

In year T, excess security returns are computed by subtracting the daily control portfolio returns from the daily security returns. The excess returns in year T of the firms in the bottom 10% of the ranking are combined with equal weights to form the excess returns of the smallest market value portfolio, MV1. Similarly, the excess returns of firms in the other deciles are combined to create excess returns of the other nine market portfolios, MV2 through MV10. Under the null hypothesis, the ten market portfolios should possess identical average excess returns which should be indistinguishable from zero.

He found a difference of about 25% in the average annual return of portfolio 1 and portfolio 10, i.e., MV1 and MV10. So he rejected APT since it was not able to explain the size effect.

The above interpretation of Reinganum may not be valid. This is because the market may be acting inefficiently.

Chen (1983) commented that the most important result of APT is that only those risks that are reflected in the covariance matrix are priced, nothing else.

He computed the b(i,j)'s (the factor loadings or the FLs) using data only from odd days. Even day returns were reserved for testing purpose (i.e., they were put in a holdout sample).

He used a unique method for estimating the FLs.

The first 180 stocks in the sample (alphabetically) are selected and their sample covariance matrix was computed. The first ten factor loadings for each stock are

obtained. Five portfolios are formed using linear programming so that the resulting portfolios will balance estimation errors with other desirable properties. The first five factor loadings are produced for every stock in the sample by solving the following matrix equation.

$$\begin{bmatrix} b(1,1)*s(1)^2 & b(1,2)*s(2)^2 \ldots & b(1,k)*s(k)^2 \\ \ldots\ldots & \ldots\ldots & \ldots\ldots \\ \ldots\ldots & \ldots\ldots & \ldots\ldots \\ b(k,1)*s(1)^2 & b(k,2)*s(k)^2 . & b(k,k)*s(k)^2 \end{bmatrix} \begin{bmatrix} b(p,1) \\ \ldots \\ \ldots \\ b(p,k) \end{bmatrix} = \begin{bmatrix} \mathrm{cov}(r(p),r(1)) \\ \ldots \\ \ldots \\ \mathrm{cov}(r(p),r(k)) \end{bmatrix} \ldots (8)$$

The five portfolios are formed using Mathematical Programming. The goal of the programming is to form portfolios so that associated factor loadings possess some desirable properties which would enable solution of eqn(8) with numerical stability.

He found that all the factor risk premia are statistically significant. He concluded that firm size does not have additional explanatory power after risk is adjusted by the factor loadings.

Dhrymes, Friend, and Gultekin (1984) reexamined the empirical evidence on APT and found the Roll and Ross's tests of APT are subject to several basic limitations.

They argued that the basic methodology of analyzing small groups of securities in order to gather confirmatory or contrary evidence relative to the APT model is seriously flawed. Given the theoretic foundations put forth by Ross, one must in the empirical analysis treat all the securities symmetrically; if that is not possible because of computer software limitations, then other ways consistent with the basic requirements of the model have to be found. Analyzing small groups of securities produces results whose meaning is useless.

Second, because of the indeterminacies of factor analysis, it is not always possible to test directly whether a given factor is priced, i.e., it is not meaningful to carry out t-tests of significance on individual risk premia coefficients. The important research issue is how many factors there are and whether collectively they are priced.

Thirdly, they found that the conclusion of Roll and Ross that there are three to five factors does not appear to be robust. They found that how many factors one discovers, depends on the size of the group of securities one deals with. If one increases the number of securities, the number of factors will also in-crease.

Roll and Ross replied that if there are actually fewer than 30 "pervasive" factors generating returns, then factor analysing groups of size 30 or more is equivalent in every way except statistical power or computational cost.

Roll and Ross commented that the point that factor analysis extracts more factors with larger groups of securities or with larger time series sample size is irrelevant. There are as many factors as there are sets of assets (pairs, triplets, and so on), and that they could all be detected with a sufficiently powerful test; but almost all of them are diversifiable and thus are just as irrelevant as if the idiosyncratic term in the APT were really purely random.

Jay Shanken (1982) challenged the view that the APT is more susceptible to empirical investigation than the CAPM. He showed that the usual empirical formulation of the testable implications of the APT is inadequate, as it precludes the very expected return differentials which the theory attempts to explain. This he proved by showing that equivalent sets of securities need not confirm to the same factor model. That is, the number of factors in the respective models need not be the same. So the usual empirical formulation of the APT, when applied to these structures, may yield different and inconsistent implications concerning expected returns for given securities. The implications will be consistent if all the securities will have the same expected returns.

Shanken also argued that Roll's (1977) criticism of CAPM can also be applied against APT. If some of the factor is highly correlated with mean-variance efficient portfolio, then that factor will appear to be priced.

Shanken (1985) said that the APT restriction is, in fact, an approximate one which prices most assets well but permits arbitrarily large deviations from exact pricing on a finite set of assets. Thus it is difficult to conceive of any (finite) empirical procedure that could be used to refute the actual conclusion of the APT. So the theory is untestable in principle.

Grinblatt and Titman (1983) used the equilibrium rather than the arbitrage argument to derive the APT pricing equation for individual assets. They showed that in a finite economy where assets have non-zero idiosyncratic risk and positive aggregate supply, the APT pricing equation (3) underpredicts the expected return of the assets. But they showed that when the idiosyncratic variance is finite, and as the relative size of an asset

gets arbitrarily small, the deviation of the asset's expected return from that predicted by eqn(3) becomes arbitrarily small.

Thus we saw from the literature that

- More than two factors affect the stock returns and hence, the CAPM is not the right description of the risk-return relation-ship;
- The equilibrium version of the APT does suffer from the same limitations as does CAPM; however, the arbitrage version of APT is free from Roll's criticism as the market portfolio does not play any role in it;
- Factor analysis is one of the many ways of estimating the factor loadings and the factor scores. But it is one of the best as its robustness in detecting the true number of factors has been tested;
- The indeterminacy problem of factor analysis can be gotten rid of by dividing the sample into several groups and then testing for the groups and then testing for the equivalence of the factor structures across groups;
- Under certain loose assumptions, it has been shown by many that eqn(3) is the right description between risk and return;
- Assuming that the market is not pricing non-systematic risk, it has been found that the maximum difference between the realised returns and the returns predicted by the APT is 0.2%.

An Empirical Investigation of the Arbitrage Pricing Theory

The arbitrage Pricing Theory states that each stock's expected excess return (i.e., nominal return minus risk-free return) is determined by the stock's factor exposures. The link between the expected return and the stock factor exposures is described in equation 3 given in the previous section. The theory does not say what the factors are, how to calculate a stock's exposures to the factors, or what the weights should be in the linear combination given by Equation 3. As the literature review revealed, there are basically two ways to estimate the factor exposures and the weights. The first is the equilibrium version of the APT that links the stock returns to some macro-economic variables. Such an approach has been adopted by Chen, Roll and Ross (1986). The factors have some meaning in such a model. For example, one can take bond market

beta, unexpected changes in inflation, unexpected changes in industrial production, etc., as the factors and attempt to predict the expected returns of stocks. The second is the statistical version of the APT. Such an approach has been adopted by Roll and Ross (1980). The statistical factor analysis assumes that the factor risk exposures remain stationery over the estimation period. The logic behind this assumption is that the factor relationship is more stable than the stock relationship. For example, it is probably more valuable to know how growth stocks have done in the past than to know the past returns of a stock currently classified as a growth stock. The problem is that the stock was probably not classified as a growth stock over the earlier period. However, the factor return will give us some information on how it may perform now as a growth stock.

For the determination of the factor loading matrix, a procedure similar to that adopted by Roll and Ross (1980) was followed. The reason why the statistical version of the APT was adopted was that it saves us the task of having to define the factors. The statistical version of the APT derives the factors from the data itself and hence we don't have to define the factors before any analysis is done. Secondly, as Grinold and Kahn (1995) said, the factors may or may not be the basic driving forces of the economy. They are merely dimensions along which to analyse the risk. Hence the task of defining the factors is in fact a daunting task.

Estimation of the factor loading matrix:

First, the weekly returns (after adjusting for bonus and stock split) were computed for 173 scrips from 7.1.1984 to 6.11.93.[4] Of course, weekly returns for all the scrips were not available for the entire period. Since SPSS package deletes likewise items in case of missing data, in order to extract the maximum information from the given data, groups were formed to ensure that in each group weekly data are available for the entire period. For example, in group 1, all the 20 scrips have weekly return figures from 7.1.1984 to 6.11.1993.

Another constraint was also taken into account at the stage of group formation. Attempts were made to ensure that no two companies belonging to the same industry

[4] The author greatfully acknowledges the help given by ICFAI in having provided daily stock returns free of cost.

were in the same group. Hence, some of the groups contained less than 20 scrips. The problem in including two companies belonging to the same industry in the same group is that `industry' may turn out to be one factor in the APT equation. But we only want the systematic risk components to be reflected in the factors in the APT equation.

Then a Maximum likelihood (ML) factor analysis was performed on the stock returns. (SPSS first computes the variance- covariance matrix and then does the factor analysis on that.) This estimates the number of factors and the matrix of loadings. The following table shows the results of factor analysis performed.

Group	3 factors s.l.	4 factors s.l.	5 factors s.l.
1	.002	.008	.02
2	.0000	.0006	.01
3	.0971	.1978	.3324
4	.0055	.1812	.5456
5	.0979	.6388	.963
6	.0083	.2218	.3925
7	.0196	.0583	.0948
8	.0528	.3114	**
9	.026	.1663	.5378

Here these s.l. represents significance level. ** factor analysis could not be performed (in the sense that the differences between two likelihood functions could not be converged even after twenty iterations.)

From the above table, it is obvious that there are at least three factors and at best four factors present. Only groups 1 and 2 suggest that a fifth factor is present.

Following Roll and Ross (1980), four factors were retained. Of course, groups 5 and 8 suggest that a fourth factor is not present. But this could be more due to sampling error than due to any economic fact. After all, if the fourth factor is really absent, then it will not be priced by the market. Whether, the fourth factor is priced or not can be known at the generalised least squares regression stage itself.

The individual asset factor loading estimates from the previous step are used to explain the cross-sectional variation of individual estimated expected returns.

If we use the average returns for the entire period, the GLS cross-sectional regression is

$L_t = (B' V^{-1} B)^{-1} B' V^{-1} r(t)$ ..... (6).

$= T\, r(t)$.

Here, L_t is the factor risk premium,

V is the population variance matrix,

B is the matrix of factor loadings,

r(t) is the return of the stock in period t,

The covariance matrix of the above estimates L_t from equation (6) can be obtained by using the formula

$(1/t)\, B' v^{-1} B$ ..... (7).

It was found that in none of the groups, any of the factor is priced at a standard 5% level of significance. The results have been shown below for all the nine groups.

Factor Risk Premiums (GLS)

Group	L_1	L_2	L_3	L_4
1	.002219	.011634	-.00536	.002666
	(.032558)	(.190629)	(-.12403)	(.047833)
2	.008639	-.0065	-.0125	-.000037
	(.273496)	(-.36557)	(-1.09581)	(.002144)
3	.009013	.00329	.009684	.009842
	(.176214)	(.08075)	(.260556)	(.280782)
4	.006049	.001829	.008715	.005942
	(.177719)	(.066844)	(.446223)	(.277025)
5	.025108	-.00802	.014648	-.03127
	(.758715)	(-.23058)	(.558518)	(-1.30928)
6	.010585	.005301	.005669	.013438
	(.21422)	(.149301)	.389696)	(.146087)
7	.005563	.005297	-.00017	.005125
	(.125248)	(.101038)	(-.00416)	(.144878)
8	.004525	.001417	.011817	.001748
	(.117934)	(.022604)	(.253886)	(.035988)
9	.004525	.001417	.011817	.001748

	(.117934)	(.022604)	(.253886)	(.035988)

The bracketed figures in the above table are the student's t- statistic for the factor risk premiums. Since the variance covariance matrix is diagonal, one can test the statistical significance of the above estimates by using simple t-tests.

Roll and Ross (1980) advised against using Ordinary Least Squares Regression (OLS) while estimating the factor risk premiums because it is biased in favor of finding priced factors when, in fact, the factors are not priced. However, because of the results obtained while using GLS regression, it was decided to use OLS to see if the results are significantly different. The findings have been summarised here.

Factor Risk Premiums (OLS)

Group	L_1	L_2	L_3	L_4
1	-.01667	.03310	.095347	-.01635
	(-.16184)	(.263236)	(.682271)	(-.10391)
2	.006564	.005188	-.00069	.003133
	(8.953871)	(2.002525)	(-.22793)	(1.005322)
3	.010237	.007057	.008259	.012507
	(4.25321)	(1.142892)	(1.446818)	(1.545083)
4	.007636	.002695	.006247	-.00281
	(1.9)	(.23)	(.53)	(-.22)
5	.008535	-.00105	.009435	-.01627
	(.970017)	(-.10468)	(.846821)	(-1.13003)
6	.014202	-.00446	.000544	.003237
	(.225293)	(-.85808)	(.098396)	(.55383)
7	.007825	.007760	.001117	.003389
	(2.200145)	(2.794375)	(.244885)	(.581963)
8	.006037	-.00109	.008974	.005761
	(1.450556)	(-.29958)	(1.564737)	(1.056321)
9	.037215	.003405	-.00014	-.01363
	(4.01269)	(.0551245)	(-.53913)	(-.82189)

Here we can observe that in 5 of the 9 groups, the first factor is priced, while in only 2 of the 9 groups the second factor is priced at a 5% level of significance. In none of the groups, the third or the fourth factor is priced.

Thus we observe that even if we use OLS (where the factors are biased to appear to be priced even when they are not priced) the factors are not priced in all the groups.

Though the results seem to indicate that the factors are not priced and hence, the APT does not work in India, caution is needed before drawing such a conclusion. There can be many reasons why the findings may not be strictly correct. They are described below:

- Maximum Likelihood factor estimation technique is one of many techniques that can be used to estimate the factor risk premia and their statistical significance. As already mentioned it assumes that the factor risk exposures of the stocks remain stationery over time which need not be a valid conclusion.
- Secondly, Roll and Ross (1980) tested the APT using daily stock returns. Here, the APT was tested using weekly returns. This substantially reduced the number of data points. One may find a different result by taking daily data points.
- The first factor seems to be the market factor as except for five stocks, the remaining stocks exhibited a positive factor loading on factor 1. In fact, the first factor explained on an average, 27% of the variation of the stock returns. Hence, it is surprising to find that a factor that explains as much as 27% of the variation in stock returns is not priced.

Implications:

These findings have got some very interesting implications for the theory of financial economics. Since the factor risks are not priced, it means the market prices the non-systematic risk. Here it is not a questionof whether the model is mis-specified or not. The fact that factor analysis is able to extract four factors means that these are probably some unspecified sources of risk that are explaining the entire variation in stock returns. However, they are not priced. And that is where the fun lies. This may be because people in India do not hold diversified portfolios. Things may change now.

Recent empirical work in US have found that book-to-market and size of the stocks are able to explain most of the variation in stock returns in US. (See for example, Fama and French (1992,1993). Fama and French explained such anomalies by saying that probably book-to-market and size are some proxies for factor risk.

The empirical anomalies that have been found cannot be explained by this because the factors are not priced at all.

Finally, the most pertinent question that has to be addressed is whether we can make arbitrage profits.Following Ross' argument, we can always package the securities in such a manner that they will look identical. Here since the factors are not priced, one can construct a portfolio that will yield us the risk-free return (in this case assumed to be 12%). This will be discussed in a leter working paper.

References

1. Brown,S., and M.Weinstein, " A New Approach to Testing Asset Pricing Models: The Bilinear Paradigm", The Journal of Finance 38 (June 1983) 711-744.
2. Chen,N.F., "Some Empirical Tests of the Arbitrage Pricing Theory", The Journal of Finance (December 1983) 1393-1414.
3. Chen,N.F., and J.E.Ingersoll,jr., "Exact Pricing in Linear Factor Models with Infinitely Many assets: A Note", The Journal of Finance, (June 1983) 958-988.
4. Chow,C.E., E.Elton and M.Gruber, "On The Robustness of the Roll and Ross Arbitrage Pricing Theory", Journal of Financial and Quantitative Analysis. 29 (March 1984) 1-10.
5. Dhrymes,P., I.Friend, and N.B.Gultekin, "A Critical Reexamination of the Empirical Evidence on the APT", The Journal of Finance 39 (June 1984) 323-34-346.
6. Dybvig,P.H., and S.Ross, "Yes, The APT is Testable", The Journal of Finance (September 1985) 1173-1188.
7. Fama, E F., and Kenneth French, "The cross-section of expected stock returns", The Journal of Finance (1985). 427-465.
8. Fama, E F., and Kenneth French, "Common risk factors in the returns on stocks and bonds", Journal of Financial Economics (1993) 3-56
9. Grinblatt,M., and S.Titman, "Factor Pricing in a Finite Economy", Journal of Financial Economics, (1983) 497-507.
10. Grinold,L., and Ronald N Kahn, "Active Portfolio management"
11. Reinganum,M., "The APT: Some Empirical Results", The Journal of Finance (May 1981) 313-321.
12. Reinganum, M.R., "Misspecification of Capital Assets Pricing: Empirical Anomalies based on Earnings Yields and Market Values", Journal of Financial Economics., (9) 1981. pp 19-46.
13. Roll,R., "A Critique of the Asset Pricing Theory' Tests. PartI: on Past and Potential Testability of the Theory", Journal of Financial Economics. 4 (1977) 129-176.
14. Roll,R., and S.Ross, "An Empirical Investigation of the APT", The Journal of Finance, (December 1980) 1073-1103.
15. Roll,R., and S.Ross, "A Critical Reexamination of the Empirical Evidence on the APT: A Reply", The Journal of Finance 39 (June 1984) 347-350.
16. Roll,R., and S.A.Ross, "On the Cross-Sectional Relation Between Expected Returns and Betas", The Journal of Finance (march 1994) 101-121.
17. Ross,S.A., "The Arbitrage Theory of Capital Asset Pricing", Journal of Economic Theory. 13 (1976) 341-360.
18. Ross,S.A., "Return, Risk, and Arbitrage" in Risk and Return in Finance (I.Friend and J.Bickster Eds.), Ballinger, Cambridge,Mass 1976.
19. Ross,S.A., "Mutual Fund Separation in Financial Theory- The Separating Distributions", Journal of Financial Economics 29 (September 1978) 254-286.
20. Shanken,J., "The APT: Is It Testable?", 37 (December 1982). 1129-1140.
21. Shanken,J., "Multi-Beta CAPM or Equilibrium APT?: A Reply", The Journal of Finance (September 1985) 1189-1196.