Sankhyā: The Indian Journal of Statistics
2001, Volume 63, Series B, Pt. 1, pp. 108–121

CAPITAL ASSET PRICING MODEL WHEN DATA IS SKEWED

By DIGANTA MUKHERJEE
and
SAMIK METIA
Indian Statistical Institute, Kolkata, India

SUMMARY. Capital Asset Pricing Models (CAPM) describe how the expected return of an asset is determined in a securities market. We assume that the individual investors are risk averse. An often observed but usually ignored feature of the distribution of security returns is its skewness. To modify the usual CAPM to allow for skewness, we attempt to change the mean-variance set up and try to work with a different metric. We explore the theoretical implications of the model and study its optimality properties. We apply our methodology on Toronto stock exchange data.

1. **Introduction**

Capital asset pricing models (CAPM) evolved out of the consumer's choice problem when faced with uncertainty. These models have been a major subject of research in finance theory. It describes how the price of a claim to a future payoff is determined in the securities market. To be more specific, CAPM describes the expected rate of return of financial assets like stocks, bonds, futures, options and other securities. The literature grew out of the works of Markowitz (1952), Sharpe (1964), Lintner (1969), Mossin (1966), Black (1972) among others. These works revolve around the mean-variance model of asset choice first developed by Markowitz (1952). A preference for expected return and aversion to variance is implied by monotonicity and strict concavity of an individual's utility function. However, for arbitrary distributions and utility functions, expected utility can not be defined on just the expected returns and variances. Nevertheless, the mean-variance model of asset choice is popular because of its analytical tractability and its rich empirical implications.

Under the usual assumptions like the *law of one price, no arbitraging* and *equilibrium in the financial market*, the mean-variance model can be motivated by two

Paper received October 2000.
AMS (1991) *subject classification.* Primary 62P05; secondary 62P20, 62P25.
Key words and phrases. CAPM, portfolio frontier, median, mean absolute deviation, skewness.

technical assumptions. We assume that there are $n \geq 2$ risky assets traded in a frictionless economy, where unlimited short selling is allowed and the rates of returns on these assets have finite variance and unequal expectations. It is also assumed that the random rate of return on any asset can not be expressed as a linear combination of rates of return on other assets. That is, asset returns are linearly independent.

For arbitrary distributions of returns, the mean-variance model is supported by assuming a quadratic (expected) utility function $u(.)$ for any asset with random return $\widetilde{W}$.

$$E[u(\widetilde{W})] = E[\widetilde{W}] - \frac{b}{2}E[\widetilde{W}^2] = E[\widetilde{W}] - \frac{b}{2}[(E(\widetilde{W}))^2 + \sigma^2(\widetilde{W})], \tag{1}$$

where $E(.)$ is the expectation operator and $\sigma^2(.)$ is the variance. Under quadratic utility, the third and higher order derivatives are zero and hence an individual's asset choice is completely determined in terms of a preference relation defined over the mean and variance of expected returns. But this utility function displays undesirable properties like satiation (an increase in wealth beyond the satiation point decreases utility) and increasing absolute risk aversion (risky assets are inferior goods). Thus, conclusions based on the assumption of quadratic utility function are often counter intuitive.

For arbitrary preferences, the mean-variance model can be motivated by assuming that rates of returns on risky assets are multivariate normally distributed. The normal distribution is completely described by its mean and variance. Thus, for utility functions that are defined over a normally distributed end of period wealth, the assumption that asset returns are multivariate normally distributed implies that demand for risky assets are defined over the mean and variance of portfolio rates of return. Unfortunately, the normal distribution is unbounded from below, which is inconsistent with limited liability. For a detailed discussion on this issue, see Huang and Litzenberger (1988).

The first step to develop the CAPM is to explore the analytical relations between the mean and variance of rates of return on feasible portfolios. This relation is graphically described as the portfolio frontier. In a real life data situation, the above assumptions for justifying the mean-variance set up is difficult to justify. The data on the rates of return we have from the Toronto stock exchange (TSE) does not match with them. The histogram of the returns of the companies show that the data is very much skewed. So, the assumption of normality can not be a good approach to build up the portfolio frontier.

In this paper we attempt to modify the estimation procedure for the portfolio frontier by allowing for skewness in the data in our methodology. This can be done by either using a different metric or by using a function of higher order moments instead of just variance.

The next section describes the two approaches that we have developed; the model and the estimation algorithm. Section 3 describes the data we have used for empirical purposes and interprets the empirical findings. Section 4 provides some concluding remarks, including plausible theoretical justifications for the results.

2. The Estimation Models and Algorithm

2.1 *Background.* Before we initiate our discussion on how to modify the estimation method of the portfolio frontier, we first formally define a frontier portfolio. In the usual mean-variance set up, a portfolio is a frontier portfolio if it has the minimum variance among portfolios that has the same expected return. To make our discussion more precise, consider $n(\geq 2)$ risky assets where returns are given by the random variables $\widetilde{r_i}$, $i = 1, ..., n$. By assumption, these are linearly independent, so the covariance matrix V with elements $Cov(\widetilde{r_i}, \widetilde{r_j})$ is positive definite. A portfolio p is defined by $\mathbf{w}_p'\tilde{\mathbf{r}}$ where $\tilde{\mathbf{r}} = (\widetilde{r_1}, ..., \widetilde{r_n})$ and $\mathbf{w}_p$ is the n-vector of portfolio weights. Now, in terms of this notation, a frontier portfolio is the solution to the quadratic program

$$\min_{\mathbf{w}} \frac{1}{2}\mathbf{w}'V\mathbf{w} \quad \text{such that} \quad \mathbf{w}'\mathbf{e} = E(r_p) \quad \text{and} \quad \mathbf{w}'\mathbf{1} = 1, \tag{2}$$

where $\mathbf{e}$ is the n-vector of expected rates of return on the n risky assets, $E(r_p)$ denotes the expected rate of return desired, $\mathbf{1}$ is the n-vector of ones. The solution to this problem is given by

$$\mathbf{w}_p = \mathbf{g} + \mathbf{h}E(r_p) \tag{3}$$

where $\mathbf{g}$ and $\mathbf{h}$ are functions of V and $\mathbf{e}$.

We will now try to motivate the two approaches we have attempted in this paper from the above discussion.

2.2 *The mean absolute deviation approach.* The quadratic program (2) is essentially aiming at minimizing mean-square-deviation or what is known in the statistics literature as the L^2 norm. This is the most popular metric used in statistics. Instead of L^2 distance minimization on the returns, we propose the use of the *mean absolute deviation* or L^1 norm. This is more suitable in the presence of outliers in the data and possible non-normality in the sense that this gives us an estimate which can be theoretically shown to be more robust. The expectation operator is replaced by the median as the suitable measure of central tendency in the objective function. Now let us formally define our estimation method.

In this approach, a portfolio p is a frontier portfolio if and only if $\mathbf{w}_p$, the n-vector of portfolio weights of p, is a solution to the quadratic program

$$\min_{\mathbf{w}} \frac{1}{2}\mathbf{w}'Q\mathbf{w} \quad \text{such that} \quad \mathbf{w}'\mathbf{e} = med(r_p) \quad \text{and} \quad \mathbf{w}'\mathbf{1} = 1, \tag{4}$$

Note that, here $\mathbf{e}$ denotes the n-vector of median returns on the n risky assets, $med(r_p)$ is the expected median return on the portfolio p and Q is a suitable distance matrix. To define Q, we first define the distance measure L_s^1 we are working with. s denotes the sample version, the sample size (number of observations) being T.

$$L_s^1 = \frac{1}{T}\sum_{t=1}^{T} |\mathbf{w}'\mathbf{x}_t - \mathbf{w}'\mathbf{e}| = \frac{1}{T}\sum_{t=1}^{T} \frac{(\mathbf{w}'\mathbf{x}_t - \mathbf{w}'\mathbf{e})^2}{|\mathbf{w}'\mathbf{x}_t - \mathbf{w}'\mathbf{e}|}$$

$$= \mathbf{w}' \left(\frac{1}{T} \sum_{t=1}^{T} \frac{(\mathbf{x}_t - \mathbf{e})(\mathbf{x}_t - \mathbf{e})'}{|\mathbf{w}'\mathbf{x}_t - \mathbf{w}'\mathbf{e}|} \right) \mathbf{w} \;=\; \mathbf{w}'Q\mathbf{w},$$

where

$$Q = \frac{1}{T} \sum_{t=1}^{T} \frac{(\mathbf{x}_t - e)(\mathbf{x}_t - \mathbf{e})'}{|\mathbf{w}'\mathbf{x}_t - \mathbf{w}'\mathbf{e}|}.$$

Now, with Q as defined above, it is not possible to get an analytical minimizing expression directly. So we take recourse to an iterative procedure. We formulate a reweighting scheme in the following fashion. At the i^{th} stage of iteration, we define

$$Q^i = \frac{1}{T} \sum_{t=1}^{T} \frac{(\mathbf{x}_t - \mathbf{e})(\mathbf{x}_t - \mathbf{e})'}{|\mathbf{w}'^{(i-1)}\mathbf{x}_t - \mathbf{w}'^{(i-1)}\mathbf{e}|}$$

where $\mathbf{w}^{(i-1)}$ is the minimizing weight vector at the $(i-1)^{th}$ stage with $\mathbf{w}^0$ being the initial weight to start the iterative procedure which we describe below.

Given the weights at the i^{th} stage, we carry out the minimisation of $\mathbf{w}'Q^i\mathbf{w}$ in the usual manner subject to the constraints. Let $\mathbf{w}_p^i$ be the minimising weights at the i^{th} stage, then $\mathbf{w}_p^i$ is the solution to the following

$$min_{\mathbf{w},\gamma,\lambda} L_s^1 = \frac{1}{2}\mathbf{w}'Q^i w + \lambda(med(r_p) - \mathbf{w}'\mathbf{e}) + \gamma(1 - \mathbf{w}'\mathbf{1})$$

where λ and γ are suitable constants. We derive the first order conditions for the above minimisation:

$$\frac{\delta L_s^1}{\delta \mathbf{w}} = Q^i\mathbf{w}_p - \lambda\mathbf{e} - \gamma\mathbf{1} = 0 \quad (5)$$

$$\frac{\delta L_s^1}{\delta \lambda} = med(r_p) - \mathbf{w}_p'\mathbf{e} = 0 \quad (6)$$

$$\frac{\delta L_s^1}{\delta \mathbf{w}} = 1 - \mathbf{w}_p'\mathbf{1} = 0 \quad (7)$$

Since we are dealing with matrices which can be expressed as the sum of positive definite matrices and number of observations is large compared to the dimension of p, we can safely assume that we have nearly positive definite matrices to work with. Then first order conditions are sufficient for the globally optimal solutions assuming Q is fixed at each stage given the previous steps. From (5) we have

$$\mathbf{w}_p = \lambda(Q^i)^{-1}\mathbf{e} + \gamma(Q^i)^{-1}\mathbf{1}. \quad (8)$$

Pre-multiplying by $\mathbf{e}'$ and using (6) we get

$$med(r_p) = \lambda(\mathbf{e}'(Q^i)^{-1}\mathbf{e}) + \gamma(\mathbf{e}'(Q^i)^{-1}\mathbf{1}). \quad (9)$$

Again pre-multiplying (8) by $\mathbf{1}'$ and using (7) we get

$$\lambda(\mathbf{1}'(Q^i)^{-1}\mathbf{e}) + \gamma(\mathbf{1}'(Q^i)^{-1}\mathbf{1}) = 1. \quad (10)$$

From (9) and (10) we obtain

$$\lambda = (C[med(r_p)] - A)/D$$
$$\gamma = (B - A[med(r_p)])/D$$

where $A = \mathbf{1}'(Q^i)^{-1}\mathbf{e}$, $B = \mathbf{e}'(Q^i)^{-1}\mathbf{e}$, $C = \mathbf{1}'(Q^i)^{-1}\mathbf{1}$ and $D = BC - A^2$.

Substituting λ and γ in (8), we finally obtain

$$\mathbf{w}_p^i = \mathbf{g} + \mathbf{h}[med(r_p)] \tag{11}$$

where $\mathbf{g} = (B((Q^i)^{-1}\mathbf{1}) - A((Q^i)^{-1}\mathbf{e}))/D$ and $\mathbf{h} = (C((Q^i)^{-1}\mathbf{e}) - A((Q^i)^{-1}\mathbf{1}))/D$. Using these new optimal weights $\mathbf{w}_p^i$ we recompute Q^{i+1} and continue in an iterative fashion.

The stopping rule: At each step we calculate $S^i = |\,||\mathbf{w}_p^i|| - ||\mathbf{w}_p^{i+1}||\,|$. We stop after j steps if $S^j < \epsilon$ where $\epsilon = 0.001$, a small prefixed quantity. The norm used here to measure the change in $\mathbf{w}_p$ over successive stages is the usual L^2 norm on vectors.

Choice of initial value $\mathbf{w}^0$: We start with $\mathbf{w}^0 = argmin(\mathbf{w}'V\mathbf{w})$ under the usual constraints, where V is the usual sample covariance matrix but with the sample mean replaced by the sample median of the rates of returns. That is, we initially start with a weight vector which is close to the one obtained from the usual mean-variance approach and then go on to modify this to take into consideration the possible effects of skewness and non-normality.

Once the procedure converges, we plot $\mathbf{w}'_P Q\mathbf{w}_p$ against $med(r_p)$ to get the portfolio frontier for this expected value of the median. We then repeat this procedure for a suitable range of values for the median to generate the entire efficient part of the portfolio frontier.

2.3 *Incorporating skewness with variance.* In this second approach, instead of considering a different metric, we incorporate the standard measure of skewness in the minimizable objective instead of considering only variance. This is done by using a linear combination of the measures of skewness and variance which we call a proxy measure of dispersion L_s. We define the sample version of the proxy measure as

$$L_s = \gamma\frac{1}{T}\sum_{t=1}^{T}(\mathbf{w}'(\mathbf{x}_t - \mathbf{e}))^2 + (1-\gamma)\left\{\frac{1}{T}\sum_{t=1}^{T}(\mathbf{w}'(\mathbf{x}_t - \mathbf{e}))^3\right\}^2$$

where $0 \le \gamma \le 1$, $\mathbf{w}'\mathbf{e} = \sum_{i=1}^{n} w_i[med(x_i)] = med(r_p) = q_p$ (say) and $\sum_i w_i = 1$. Note that we have taken the square of the third moment as our indicator of skewness to take care of a possible negative sign cancelling out with variance.

The above L_s is to be minimised with respect to $\mathbf{w}$ under the constraints as before. So we form the Lagrangian as follows,

$$min_{\mathbf{w},\lambda_1,\lambda_2} L = L_s + \lambda_1\left(\sum w_i - 1\right) + \lambda_2\left(\sum w_i x_i - q_p\right).$$

Here we will be using a Newton-Raphson type iterative procedure for a multi-parameter set up. We calculate the following quantities in the iterative steps.

For $1 \leq j \leq m$,

$$\begin{aligned} S_j &= \frac{\delta L}{\delta w_j} = 2\gamma \frac{1}{T}\sum_{t=1}^{T}(\mathbf{w}'(\mathbf{x}_t - \mathbf{e}))(x_{jt} - \widetilde{x}_j) \\ &\quad + 2(1-\gamma)\left\{\frac{1}{T}\sum_{t=1}^{T}(\mathbf{w}'(\mathbf{x}_t - \mathbf{e}))^3\right\} \\ &\quad \cdot \left\{\frac{3}{T}\sum_{t=1}^{T}(\mathbf{w}'(\mathbf{x}_t - \mathbf{e}))^2(x_{jt} - \widetilde{x}_j) + \lambda_1 + \lambda_2\widetilde{x}_j\right\}. \\ S_{m+1} &= \frac{\delta L}{\delta \lambda_1} = \sum_{i=1}^{m} w_i - 1, \\ S_{m+2} &= \frac{\delta L}{\delta \lambda_2} = \sum_{i=1}^{m} w_i\widetilde{x}_i - q_p. \end{aligned}$$

Define $\mathbf{S} = (S_1, S_2, ..., S_{m+1}, S_{m+2})'$.

For $1 \leq j, k \leq m$, let

$$\begin{aligned} I_{jk} &= \frac{\delta^2 L}{\delta w_j \delta w_k} = 2\gamma\frac{1}{T}\sum_{t=1}^{T}(x_{kt} - \widetilde{x_k})(x_{jt} - \widetilde{x}_j) \\ &\quad + \frac{6(1-\gamma)}{T}\left\{\frac{3}{T}\sum_{t=1}^{T}(\mathbf{w}'(\mathbf{x}_t - \mathbf{e}))^2(x_{kt} - \widetilde{x_k})\right\}\sum_{t=1}^{T}(\mathbf{w}'(\mathbf{x}_t - \mathbf{e}))^2(x_{jt} - \widetilde{x}_j) \\ &\quad + \frac{6(1-\gamma)}{T}\left\{\frac{1}{T}\sum_{t=1}^{T}(\mathbf{w}'(\mathbf{x}_t - \mathbf{e}))^3\right\} 2\sum_{t=1}^{T}(\mathbf{w}'(\mathbf{x}_t - \mathbf{e}))(x_{kt} - \widetilde{x_k})(x_{jt} - \widetilde{x}_j). \end{aligned}$$

$$\begin{aligned} I_{m+1,j} &= I_{j,m+1} = \frac{\delta^2 L}{\delta \lambda_1 \delta w_j} = 1, \quad j = 1, 2, ..., m. \\ I_{m+2,j} &= I_{j,m+2} = \frac{\delta^2 L}{\delta \lambda_2 \delta w_j} = \widetilde{x}_j, \quad j = 1, 2, ..., m. \\ I_{m+1,m+1} &= I_{m+1,m+2} = I_{m+2,m+2} = 0. \\ \mathbf{I}_{(m+2)\times(m+2)} &= ((I_{jk})). \end{aligned}$$

Define the i^{th} stage value by

$$\mathbf{w}*^{(i)} = \begin{pmatrix} \mathbf{w}^i \\ \lambda_1^i \\ \lambda_2^i \end{pmatrix}.$$

Using Newton-Raphson method, we get the $(i+1)^{th}$ stage iterate as

$$\mathbf{w}*^{(i+1)} = \mathbf{w}*^{(i)} - \mathbf{I}^{-1}|_{\mathbf{w}*^{(i)}}\mathbf{S}^i$$

where $\mathbf{I}^{-1}|_{\mathbf{w}*^{(i)}}$ and $\mathbf{S}^i$ are recomputed at every stage.

Stopping rule: We continue upto the j^{th} step where we have

$$\left| ||\mathbf{w}*^{(i+1)}|| - ||\mathbf{w}*^{(i)}|| \right| < \epsilon,$$

where ϵ is a prefixed small positive quantity.

After convergence, we need to check all the eigenvalues of the $\mathbf{I}$ matrix at that stage of iteration. If all the eigenvalues are positive, then only we can say that a local minima has been reached for the proxy measure using the weights thus obtained, if not the global minimum. To illustrate our method, we chose three alternative values for γ, namely 0.25, 0.5 and 0.75, to check the effect over the entire range of values for γ. Basically, these different weights give different levels of importance to the skewness in the data which is revealed through the second term of the proxy measure. Smaller values of γ imply greater weightage to the skewness factor.

Choice of initial value: As we are using the Newton-Raphson method, the choice of initial value is of great importance to us in terms of convergence. For some choices, the procedure may not converge at all. The best possible choice would have been to start with uniform initial weights. But that is not possible here as we have a set of constraints which won't allow uniform weights. So we tried an intuitively appealing set of initial weights which is close to uniform. We consider

$$\begin{aligned} w_i^0 &= \frac{1}{m}, i = 1, 2, ..., m-2 \\ w_{m-1}^0 &= \frac{q_p - \frac{1}{m}\sum \widetilde{x_j} - \frac{2}{m}\widetilde{x_m}}{\widetilde{x_m} - \widetilde{x_{m-1}}}, \quad \widetilde{x_m} \neq \widetilde{x_{m-1}} \\ w_m^0 &= \frac{2}{m} - w_{m-1}^0 \\ \lambda_1^0 &= \lambda_2^0 \; = \; 1. \end{aligned}$$

3. **Empirical Analysis**

3.1 *The data.* The data used here for empirical purposes is from the Toronto Stock Exchange, Canada which is obtained from the internet site:
http://metalab.unc.edu/pub/archives/misc.invest/historical-data/stocks/canada
The time period for which the data was collected is 1993–1996. It includes companies from the following sectors: OIL and GAS (4 companies), METALS and MINERALS (4), FINANCIAL (2), SOFTWARE (3), CONSUMER PRODUCT (4), PAPER and FOREST (4), INDUSTRIAL PRODUCTS (3), TRAVEL and TOURISM (1), PHARMACEUTICAL (1) and REAL ESTATE (2). There were 50 companies which we considered at the start. However, we used the criteria of stochastic dominance (Risky asset A is said to stochastically dominate, in the second degree, risky asset B if A has the same expected rate of return as B and a lower variance.) to bring this down to 28. We finally worked with these 28 companies.

To calculate the rate of return on day $(n+1)$ we used

$$RoR = \frac{\text{opening stock at } (n+1)^{th} \text{ day} - \text{opening stock at } n^{th} \text{ day}}{\text{opening stock at } n^{th} \text{ day}}$$

3.2 *Empirical results for the mean absolute deviation approach.* The estimated portfolio frontier for the real data and the simulated multivariate normal data are shown in Figures 1 and 2. From Figure 1, it is seen that the efficient part of the portfolio frontier is almost linear. In Section 4, we provide a plausible theoretical justification for this. Our intuition is strengthened when we look at Figure 2. Even with the simulated multivariate normal data, the portfolio frontier looks like a combination of two straight lines.

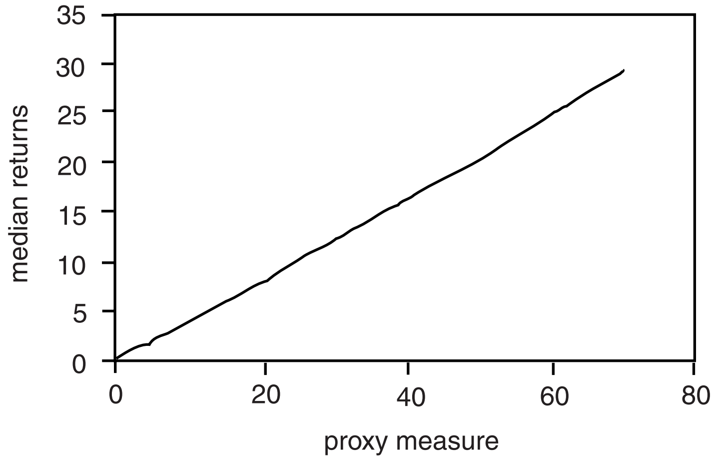

Figure 1. Portfolio frontier with real data sets (using L1)

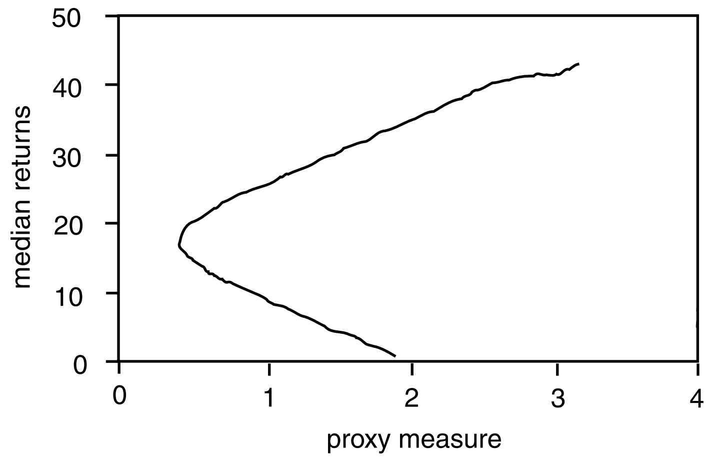

Figure 2. Portfolio frontier with simulated mvn data (using L1)

3.3 *Empirical results for the proxy measure.* We first carried out the estimation algorithm to build the portfolio frontier for a simulated multivariate normal data. We chose three values for γ, namely 0.25, 0.5 and 0.75 for our study. The result we got was close to the parabolic form in each case. But as skewness was incorporated here, and $\gamma = 0.25$ gives highest weightage to it, we got disturbances in the normal parabolic form in that case (Figure 3). Such disturbances were lesser for the higher γ values (as in Figure 4 for $\gamma = 0.75$).

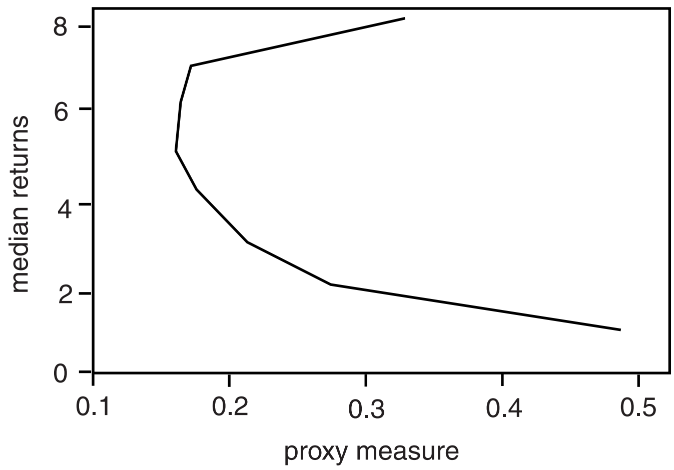

Figure 3. Portfolio frontier with mvn data with $\gamma = 0.25$ (using proxy measure)

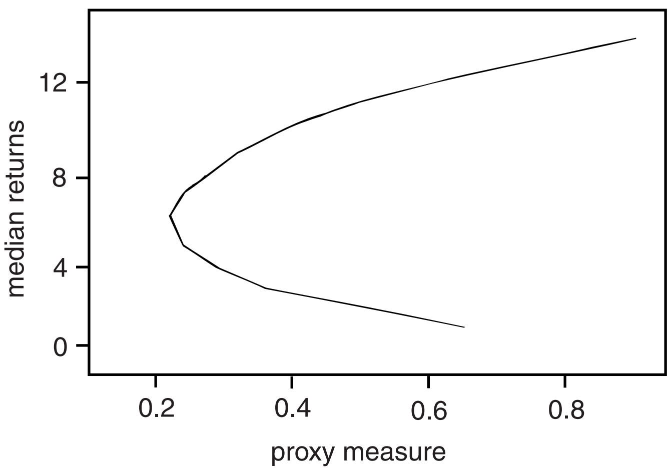

Figure 4. Portfolio frontier with mvn data with $\gamma = 0.75$ (using proxy measure)

Next we looked at the TSE data set. As mentioned earlier, we worked with the stochastically undominated returns only. Since we did not make any distributional assumptions on the rates of return, we needed to impose some restriction on the underlying utility function. We considered the *Von Neumann-Morgenstern* utility function

$$u(\mathbf{w}'\widetilde{\mathbf{x}}) = 1 - \mathbf{e}^{-\mathbf{w}'\widetilde{\mathbf{x}}+\lambda L}$$

where L is the proxy measure defined earlier.

Now, for different values of γ chosen earlier, we get an almost parabolic form of the efficient part of the portfolio (Figures 5–7). Note that we do not observe the inefficient part of the frontier, since the investors are risk averse. All these look similar to the convex form which we obtained in the usual variance minimizing approach. Since the chosen utility function is parametrized by the median and the proxy measure only, we can expect that in the limit, any portfolio on the frontier will have the weight vector of the form

$$\mathbf{w}_p = \alpha + \beta[med(r_p)],$$

though the constants α and β are different from those obtained in the usual mean variance approach. In such a case, the market portfolio, which will be a convex combination of the other optimal portfolios, will also be on the frontier.

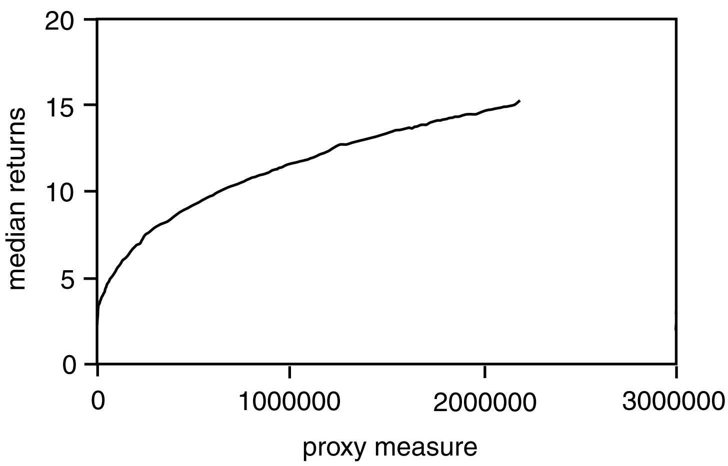

Figure 5. Portfolio frontier with $\gamma = 0.5$ (using proxy measure)

4. Interpretation of the Empirical Results and Conclusion

To interpret the results we have obtained using the approaches discussed above, we first look at the usual results which follows from the assumption of multivariate normality of the underlying distribution. In the usual CAPM, for any portfolio p which is not the minimum variance portfolio, we have a nice expression for the zero covariance counterpart. Geometrically, its the point of intersection of the portfolio frontier with the horizontal line drawn from the point where the tangent to the frontier at p intersects the vertical axis, i.e. the variance axis. Using the algebraic expression

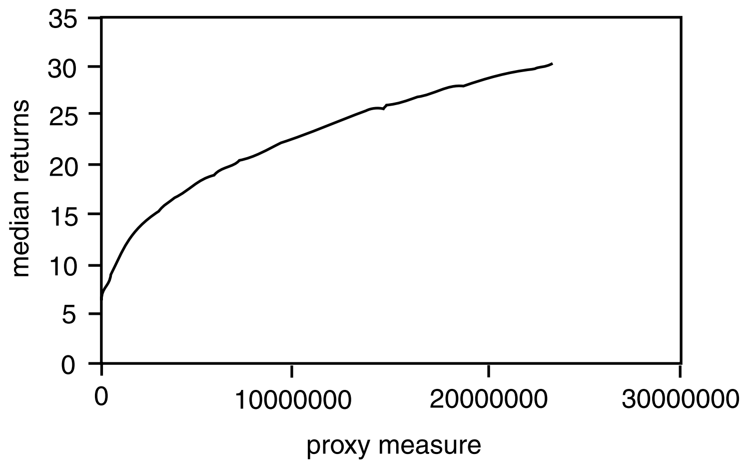

Figure 6. Portfolio frontier with $\gamma = 0.25$ (using proxy measure)

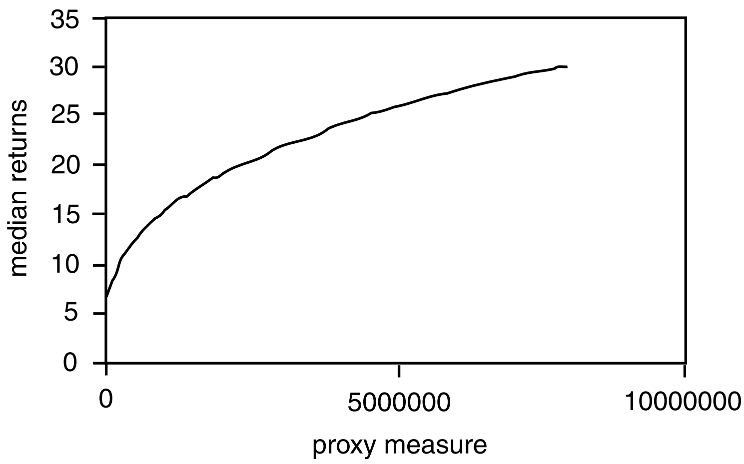

Figure 7. Portfolio frontier with $\gamma = 0.75$ (using proxy measure)

for this zero covariance portfolio $zc(p)$ of p, we can express any portfolio as a linear combination of p and $zc(p)$ with some zero mean error. More precisely, we obtain

$$r_q = b_0 + b_1 r_{zc(p)} + b_2 r_p + \epsilon_q$$

where $Cov(r_{zc(p)}, \epsilon_q) = E(\epsilon_q) = 0$. The portfolios p and $zc(p)$ are called the separating portfolios and this property is termed two funds separation. These separating portfolios are always on the portfolio frontier and for the usual CAPM setup the market portfolio of returns and its orthogonal counterpart are taken as the separating portfolios.

We have tried to develop our alternative algorithms in the above line of approach. the results we obtain are quite interesting.

4.1 *The mean absolute deviation results.* Here, we get almost linear portfolio function when we work with the TSE data. We try to explain this linearity when we consider two returns only. We have

$$w_1\widetilde{x_1} + w_2\widetilde{x_2} = q, \quad w_2 = 1 - w_1 \Rightarrow w_1 = \frac{|q - \widetilde{x_2}|}{|\widetilde{x_1} - \widetilde{x_2}|}$$

where we have assumed that $\widetilde{x_1} \neq \widetilde{x_2}$. This implies

$$E|\mathbf{w}'\mathbf{x} - \mathbf{w}'\widetilde{\mathbf{x}}| = E\left|(q - \widetilde{x_2})\left(\frac{x_1 - x_2}{\widetilde{x_1} - \widetilde{x_2}} - 1\right) + (x_2 - \widetilde{x_2})\right|$$

which is linear in q.

This argument can be generalized to the more than two returns case. This seems to give an explanation for the linearity observed here. Even though the linear portfolio frontier leads to the closure property and a separation result as discussed above, we do not get any closed form expressions for the separating portfolios. The method for finding out the zero covariance portfolio is not transparent. Ofcourse this argument entails that we must consider an utility function such that the indifference curves touch the linear frontier at unique points. So we need to have indifference curves with a strict curvature to do suitable analysis in this situation.

4.2 *The proxy measure results.* We refer to the Figures 3–7. Here, we get the almost parabolic form as in the usual mean variance approach. So we would like to get a structure akin to the two fund separation result obtained in the classical approach. (For a detailed discussion on the empirical methods involving the estimation of CAPM, see Rao and Krishnaiah, 1994.)

To get towards the two fund separation, we tried to find the zero covariance portfolio in this case. But, we did not get any nice algebraic form. We obtain

$$\begin{aligned} L &= \gamma\frac{1}{n}\sum_{t=1}^{T}(\mathbf{w}'(\mathbf{x}_t - \mathbf{e}))^2 + (1-\gamma)\left[\frac{1}{n}\sum_{t=1}^{T}(\mathbf{w}'(\mathbf{x}_t - \mathbf{e}))^3\right]^2 \\ &= \gamma\frac{1}{n}\sum_{t=1}^{T}\mathbf{w}'V_t\mathbf{w} + (1-\gamma)\left[\frac{1}{n^2}\left\{\sum_{t=1}^{T}(\mathbf{w}'(\mathbf{x}_t - \mathbf{e}))^6\right.\right. \\ &\quad \left.\left. + \sum_{t\neq}\sum_{u}(\mathbf{w}'(\mathbf{x}_t - \mathbf{e}))^3(\mathbf{w}'(x_u - \mathbf{e}))^3\right\}\right] \end{aligned}$$

where

$$\begin{aligned} V_t &= (\mathbf{x}_t - \mathbf{e})(\mathbf{x}_t - \mathbf{e})'. \\ &= \gamma\frac{1}{n}\sum_{t=1}^{T}\mathbf{w}'V_t\mathbf{w} + (1-\gamma)\frac{1}{n^2}\left[\sum_{t=1}^{T}\{\mathbf{w}'(\mathbf{x}_t - \mathbf{e})(\mathbf{x}_t - \mathbf{e})'\mathbf{w}\}\right. \\ &\quad \cdot\{\mathbf{w}'(\mathbf{x}_t - \mathbf{e})(\mathbf{x}_t - \mathbf{e})'\mathbf{w}\}\{\mathbf{w}'(\mathbf{x}_t - \mathbf{e})(\mathbf{x}_t - \mathbf{e})'\mathbf{w}\} \\ &\quad \left. + \sum_{t\neq}\sum_{u}\mathbf{w}'(\mathbf{x}_t - \mathbf{e})(\mathbf{x}_t - \mathbf{e})'\mathbf{w}\mathbf{w}'(\mathbf{x}_t - \mathbf{e})(x_u - \mathbf{e})'\mathbf{w}\mathbf{w}'(\mathbf{x}_u - \mathbf{e})(\mathbf{x}_u - \mathbf{e})'\mathbf{w}\right]. \end{aligned}$$

Let $\widetilde{\mathbf{w}_p}$ be the optimal weight for the return vector $\widetilde{r_p}$ on portfolio frontier. Then,

$$L(\widetilde{r_p}) = \gamma\frac{1}{n}\sum_{t=1}^{T}\mathbf{w}_p' V_t \mathbf{w}_p + (1-\gamma)\frac{1}{n^2}\left[\sum_{t=1}^{T}(\mathbf{w}_p' V_t \mathbf{w}_p)^3 + \sum_{t\neq}\sum_{u}(\mathbf{w}_p' V_t \mathbf{w}_p)(\mathbf{w}_p' V_{t,u}\mathbf{w}_p)(\mathbf{w}_p' V_u \mathbf{w}_p)\right].$$

For this approach $L(\widetilde{r_p})$ is the counterpart of $Var(\widetilde{r_p})$. We define the counterpart of the covariance term as

$$L(\widetilde{r_p},\widetilde{r_q}) = \gamma\frac{1}{n}\sum_{t=1}^{T}\mathbf{w}_p' V_t w_q + (1-\gamma)\frac{1}{n^2}\left[\sum_{t=1}^{T}(\mathbf{w}_p' V_t \mathbf{w}_p)(\mathbf{w}_p' V_t w_q)(w_q' V_t w_q) + \sum_{t\neq}\sum_{u}(\mathbf{w}_p' V_t \mathbf{w}_p)(\mathbf{w}_p' V_{t,u} w_q)(w_q' V_u w_q)\right].$$

The only striking feature here is the $V_{t,u}$ matrix, which tries to capture the skewness within the data. It compares two time points instead of comparing a point with the median and retaining the sign with it. But from the above expressions it is not easy to find out a closed form solution for the zero covariance portfolio for r_p. However, it is clear that it would depend on the relation between two time points. So, even if we try to build up an equation as in the CAPM, we may approach a multifactor model situation[1], instead of a single factor one as in the usual CAPM. Since the above equations were not analytically tractable, we suggested scoring method for finding solutions.

The form of $V_{t,u}$ is indeed quite interesting. The absence of such a term in the usual objective function seems to suggest that the classical approach imposes some sort of symmetry on the data and the estimates. If indeed the data were symmetric, then $V_{t,u}$ should be insignificant but not otherwise. As we have seen real life data to exhibit significant asymmetry, this $V_{t,u}$ term seems to have an implicit importance towards understanding the nature of the portfolio frontier.

4.3 *Concluding remarks.* In this paper, we have tried to explore the theoretical and empirical implications of two alternative methods of estimating the portfolio frontier as a first step towards building an alternative CAPM. As the methods suggested here do not yield tractable algebraic forms, we do not get closed form solutions as in the usual approach. The method is also quite computer intensive. These methods do not give us the elegant form in which one is used to seeing the

[1] In the multifactor model we have the relations in the following form:

$$E(r_t|Z_{t-1}) = E(f_t|Z_{t-1})(E(u'_{ft}u_{ft}|Z_{t-1}))^{-1}E(u'_{ft}u_t|Z_{t-1})$$

where r is a row vector of n asset returns, f is the vector of factor realisations, u_f is the vector of innovations in conditional mean of returns. The information set at time t is the data upto time point $t-1$, denoted by Z_{t-1}. In fact, the first term represents the conditional expectation of factor realisations. The second term is inverse of conditional covariance matrix and the third term is the conditional covariance of asset returns with the factors.

CAPM. But, despite all these technical problems, we find this exercise to be worthwhile as real life data sets on returns show significant departure from normality, rendering the usual mean variance approach quite suspect. One needs to drop the restrictive assumptions and look at the distribution of returns with a flexible point of view. In this context, the use of the L_1 norm or incorporating skewness is indeed useful.

Acknowledgement. The authors are grateful to Prof. Dipankar Coondoo and Dr. Snigdhanshu Bhushan Chatterjee for their illuminating comments. The general disclaimer applies.

References

Black, F. (1972). Capital market equilibrium with restricted borrowing, *Journal of Business*, **45**, 444–454.

Huang, C. and Litzenberger, R.H. (1988). *Foundations for Financial Economics*. North Holland, New York.

Lintner, J. (1969). The aggregation of inverstor's diverse judgements and preferences in purely competitive markets, *Journal of Financial and Quantitative Analysis*, **4**, 346–382.

Markowitz, H. (1952). Portfolio selection, *Journal of Finance*, **7**, 77–91.

Mossin, J. (1966). Equilibrium in a capital asset market, *Econometrica*, **35**, 768–783.

Rao, C.R. and Krishnaiah, P.R. (1994). *Handbook of Statistics*, Vol. **14**. North Holland, New York.

Sharpe, W. (1964). Capital asset prices: A theory of capital market equilibrium under conditions of risk, *Journal of Finance*, **19**, 425–442.

Diganta Mukherjee and Samik Metia
Economic Research Unit
Indian Statistical Institute
203 B.T.Road, Kolkata 700 035, India
E-mail: diganta@isical.ac.in/ digantam@hotmail.com