GS Cloete. PJ de Jongh and T de Wet*

Combining Vasicek and Robust Estimators for forecasting systematic risk

1. INTRODUCTION

The problem of estimating and forecasting systematic risk, or the so-called beta parameter in the market model, is well-known and has been studied by several authors (see e.g. Lam 1999, Lally 1998, Bowie and Bradfield 1998, Boabang 1996, Draper and Paudyal 1995, Murray 1995 and Bartholdy and Riding 1994). The classical estimator for beta is the well-known ordinary least squares (OLS) estimator, but several authors have shown that this estimator suffers from several deficiencies, e.g. it has a mean reversion tendency, is inefficient when return distributions are non-normal, and has significant bias problems when shares are thinly traded. Several alternatives to OLS have been proposed in the literature. Amongst others, Vasicek (1973) and Blume (1973) proposed estimators to improve the mean reversion tendency of beta, Chan and Lakonishok (1992) proposed robust estimators to ensure more efficient estimation of beta, and Scholes and Williams (1977) proposed estimators to deal with the bias problem when shares are infrequently traded. A host of empirical studies have been carried out in order to evaluate the performance of the estimators under various conditions (see e.g. the recent studies by Draper and Paudyal 1995, Murray 1995, Boabang 1996, and Lally 1998). Of the above-mentioned estimators, the Vasicek-estimator and the robust estimators seem to perform well over a wide range of empirical studies.

In this paper we will base the so-called Vasicek-estimator (see Vasicek 1973) on the class of L-estimators and will evaluate its performance empirically, using data from the Johannesburg Stock Exchange (JSE). By doing this we combine the properties of the Vasicek-estimator with that of robust estimators and show that, in terms of stability, the new estimators perform much better than OLS and better or as good as some of the other popular estimators. In order to combine the Vasicek-estimator with the class of L-estimators, good scale estimators are needed for the various L-estimators considered. We define suitable scale estimators and show their consistency.

*Respectively Market Risk, Standard Corporate and Merchant Bank, 3 Simmonds Street, Johannesburg 2001, South Africa, Centre for Business Mathematics and Informatics, Potchefstroom University for CHE, Private Bag X6001, Potchefstroom 2531, South Africa and Department of Statistics and Actuarial Science, University of Stellenbosch, Stellenbosch 7600, South Africa. The last author's research is supported by NRF Grant 2046922 and a grant from The University of Stellenbosch Research Fund (Research Committee A).
Email: bwipjdj@puknet.puk.ac.za

The paper is organised as follows. In Section 2 we discuss the model and estimators. Scale estimators for the L-estimators under study are defined in Section 3. In Section 4, using data from the JSE, the performance of the new estimators are compared with the least squares estimator and with the Vasicek-estimator based on OLS. Some concluding remarks and ideas for future research are given in the last section.

2. MODEL AND ESTIMATORS

The market model is given by

$$R_{it} = \alpha_i + \beta_i R_{mt} + e_{it} \quad ; i=1,\ldots,N \quad t=1,\ldots,T \qquad \ldots (1)$$

where:
R_{it} is the excess return of stock *i* in period t, α_i the unique expected return of security i,

β_i the sensitivity of stock i to market movements, R_{mt} the excess return on the market in period t, and

e_{it} the unique risky return of security i in period t, assumed to have mean zero and variance σ^2_{ei}.

It is further assumed that the errors are independently and identically distributed with respect to time according to some unknown distribution F and that the errors and the market returns are uncorrelated. For notational convenience we will suppress the potential dependence of F on i. Note that excess return refers to the return over and above the risk-free rate of return. The market model is motivated by the empirical fact that most asset returns seem to depend on the general movement of the market. The model postulates that the return on a particular asset is linearly related to the market movement plus an unexplained random component.

Under the market model it is well-known that the total risk can be written as

$$Var(R_{it}) = \beta_i^2\sigma_m^2 + \sigma_{ie}^2 \quad ,$$

where $\beta_i^2\sigma_m^2$ is the systematic risk and σ_{ie}^2 the specific risk of the i-th share. Note that beta is then a measure of the systematic risk of a particular share. Beta has several other important roles, amongst others, it can be used to estimate the covariance structure between assets, an essential input to Markowitz mean-variance portfolio optimisation. The parameter beta also plays an important role in simple portfolio optimisation procedures. For example, it is shown in Elton and Gruber (1995), that a simple ranking device based on beta leads to the optimal mean-variance portfolio on the mean-variance efficient frontier. Also, under the CAPM assumptions, beta is used to establish the fair price of an asset. There are a number of examples of the use of the above-mentioned estimators in practice and many institutions make use of so-called beta information services. For example Value Line provides Blume betas, the London Business School and Merrill Lynch provide Vasicek betas. In SA, the Financial Risk Service of the University of Cape Town provides a similar service. Clearly beta plays an important role and finding more stable estimators for beta remains an important research topic.

As mentioned before, Vasicek (1973) introduced a popular estimator for beta. Let $\bar{\beta}_1$ be the average beta across the sample of stocks in the historical period, and let β_{i1} be the OLS estimate of security i's beta measured in period one, then the Vasicek procedure involves taking a weighted average of average beta ($\bar{\beta}_1$) and the historical beta for security i (β_{i1}). Let $\sigma^2_{\bar{\beta}_1}$ denote the variance of the distribution of the historical estimates of beta over the sample of stocks. Note that this give us a measure of the uncertainty in estimating the average beta. Let $\sigma^2_{\beta_{i1}}$ denote the variance of the estimate of beta for security i measured in time period 1. Clearly, this is a measure of the uncertainty associated with the estimate of the individual securities' beta. Vasicek then suggested weights of

$$\frac{\sigma^2_{\bar{\beta}_1}}{\sigma^2_{\bar{\beta}_1}+\sigma^2_{\beta_{i1}}} \text{ for } \beta_{i1} \text{ and } \frac{\sigma^2_{\beta_{i1}}}{\sigma^2_{\bar{\beta}_1}+\sigma^2_{\beta_{i1}}} \text{ for } \bar{\beta}_1,$$

that add to 1. The forecast of beta for security i is thus

$$\beta_{i2} = \frac{\sigma^2_{\beta_{i1}}}{\sigma^2_{\bar{\beta}_1}+\sigma^2_{\beta_{i1}}}\bar{\beta}_1 + \frac{\sigma^2_{\bar{\beta}_1}}{\sigma^2_{\bar{\beta}_1}+\sigma^2_{\beta_{i1}}}\beta_{i1} \qquad \ldots (2)$$

where β_{i2} denotes the forecast. When the uncertainty about the estimate of the individual beta is high compared to the uncertainty in the average beta estimate, the beta forecast is adjusted strongly towards the average beta. On the other hand, when the uncertainty about the estimate of the individual beta is small compared to the uncertainty in the estimate of the average beta, then the beta forecast is adjusted strongly towards the individual beta. This means that high beta stocks will have their betas lowered by a bigger percentage of the distance from the average beta for the sample than low beta stocks will have their betas raised. Hence, the estimate of the average future beta will tend to be lower than the average beta in the sample of stocks over which beta is estimated. Unless there is reason to believe that the betas will continually decrease over time, the estimate of beta can be further improved by adjusting all betas upwards so that they have the same mean as they had in the historical period. Vasicek has shown that this is a Bayesian estimation technique, and it can also be seen as a so-called shrinkage estimator.

Several authors have suggested that partitioning is desirable when using the Vasicek-estimator. Partitioning involves classifying shares into different categories or groups and then basing the estimation of parameters for each share on the shares in the particular group, rather than the total population of shares considered. Lally (1998) showed that partitioning (to a larger or smaller degree) is indeed desirable. Traditionally partitioning has been done on an industry basis as suggested by Vasicek (1973). Partitioning dramatically reduces the absolute error in beta estimation, and eliminates the tendency for stocks in low beta groups to be overestimated, and those in high beta groups to be underestimated. In this study we will partition on the base of a classification based on liquidity considerations.

Clearly, it can be seen from (2) that β_{i1} can be replaced by any other estimator, provided that an appropriate estimator for $\sigma^2_{\beta_{i1}}$ can be found. In this paper we will propose members of the class of L-estimators as alternatives to OLS, the estimator on which the Vasicek-estimator is frequently based.

3. L-ESTIMATORS AND SCALE ESTIMATORS

We will consider three robust estimators, namely the L_1-estimator, the Koenker-Bassett regression trimmed mean estimator (see e.g. Ruppert and Carroll 1980) and the Welsh regression trimmed mean estimator (see Welsh 1987), and three bounded-influence versions of the same estimators. These estimators have been defined and discussed in detail in de Jongh et al. (1988) and will not be defined here because the implementation is identical. However, in order to define appropriate scale estimators for the various estimators we need some theory and have to consider the asymptotic distribution of the estimators. It can be shown that, under certain regularity conditions, all the robust and bounded-influence estimators asymptotically have distributions of the following form:

$$\sqrt{T}\left(\underline{\hat{\beta}}-\underline{\beta}\right)\xrightarrow{D}N\left(\underline{0},\lambda Q^{-1}\right) \qquad \ldots (3)$$

where Q is a limiting form depending on the design matrix in the usual regression setting. For the robust estimators, we have the usual $Q=\lim_{T\to\infty} X'X$, with X the design matrix. In the case of the bounded influence estimators, Q is slightly more complicated, depending not only on the design matrix, but also on a pre-selected weight matrix (see de Jongh et al. 1988 for details). In both cases, given the design matrix and the weight matrix, it is straightforward to approximate Q for all the robust and bounded-influence estimators. What remains to be done is to find estimators for λ. In the case of the L_1-estimator, $\lambda = \left(\frac{1}{2f\left(F^{-1}(0.5)\right)}\right)^2$, the so-called sparsity function, may be estimated (see Koenker and Machado 1999) by

$$\begin{aligned}\hat{\lambda} &= \frac{1}{4}\left(\hat{s}(0.5)\right)^2 \\ &= \frac{1}{4}\left(\frac{\underline{\bar{x}}'\hat{\underline{\beta}}_{RQ}(0.5+h_r) - \underline{\bar{x}}'\hat{\underline{\beta}}_{RQ}(0.5-h_r)}{2h_r}\right)^2\end{aligned} \quad \ldots (4)$$

with $h_r = 0.971558\,(T)^{-\frac{1}{3}}$ (assuming α=0.25 in their equation) and $\underline{\bar{x}}' = [\,1, \frac{1}{T}\sum_{t=1}^{T} R_{mt}]$ where T is the number of observations. Koenker and Machado (1999) show that $\hat{\lambda}$ is a consistent estimator for λ. Here $\hat{\underline{\beta}}_{RQ}(\alpha)$ is the α-th regression quantile as defined by Koenker and Bassett (1978). A natural estimator for λ in the case of the bounded-influence L_1-estimator will be

$$\begin{aligned}\hat{\lambda} &= \frac{1}{4}\left(\hat{s}(0.5)\right)^2 \\ &= \frac{1}{4}\left(\frac{\underline{\bar{x}}'\hat{\underline{\beta}}^{(w)}_{RQ}(0.5+h_r) - \underline{\bar{x}}'\hat{\underline{\beta}}^{(w)}_{RQ}(0.5-h_r)}{2h_r}\right)^2\end{aligned} \quad \ldots (5)$$

with $h_r = T^{-\frac{1}{3}}(0.971558)$, where $\hat{\underline{\beta}}^{(w)}_{RQ}(\alpha)$ is the α-th bounded-influence regression quantile as defined by de Jongh et al. (1988). Using the results in Koenker and Machado (1999), it easily follows that $\hat{\lambda}$, based on $\hat{\underline{\beta}}^{(w)}_{RQ}(\alpha)$, is a consistent estimator for λ.

In the case of the robust and bounded influence trimmed mean estimators, λ is the asymptotic variance of the trimmed mean in the location case, i.e.

$$\sigma^2(\alpha,F) = (1-2\alpha)^{-2}\left\{\int_{F^{-1}(\alpha)}^{F^{-1}(1-\alpha)} e^2 dF(e) + \alpha\left(F^{-1}(\alpha) + F^{-1}(1-\alpha)\right)^2\right\} \quad \ldots (6)$$

with F the distribution function (assumed symmetric around 0) and $F^{-1}(\alpha)$ the α^{th} quantile of F. An estimator for the asymptotic variance has been defined by de Jongh and de Wet (1986). Let $r_{(i)}$ be the i^{th} ordered residual from the trimmed mean fit, then the following Winsorised type estimator may be used to estimate $\sigma^2(\alpha,F)$:

$$D^2 = (1-2\alpha)^{-2}\left\{\left(\frac{1}{(T-p)}\right)\sum_{i=[T\alpha]+1}^{T-[T\alpha]} r^2_{(i)} + \alpha\left(r^2_{([T\alpha]+1)} + r^2_{(T-[T\alpha])}\right)\right\} \quad \ldots (7)$$

with [x] denoting the largest integer contained in x. De Jongh (1985) showed that D^2 is a consistent estimator for $\sigma^2(\alpha,F)$. For the bounded-influence trimmed mean estimators D^2 can again be used as scale estimator, but now $r_{(i)}$ is the i^{th} ordered residual from the bounded influence regression trimmed mean fit. Using Theorem 3 and 4 from Ruppert and Carroll (1980), one can easily show that D^2 is a consistent estimator for $\sigma^2(\alpha,F)$.

We have established consistent scale estimators for all the robust and bounded-influence estimators that we wish to consider and are now able to investigate the performance of the estimators empirically.

4. THE EMPIRICAL STUDY

In this study the performance of the estimators is evaluated by means of an empirical study. The estimators that are studied are the L-estimators (robust and bounded-influence versions) and the Vasicek estimators based on these L-estimators. The focus will be on identifying the more stable estimators for beta. This section is organised in three sub-sections. In Section 4.1, the data used to analyse the performance of the estimators is discussed. In Section 4.2, the estimators and the performance measures used to compare them are discussed and in Section 4.3 the results of the study are given.

4.1 The data analysed

In this sub-section, we discuss the data and the data source that was used as well as how the shares were selected and classified in terms of a liquidity rating. Lastly, the return calculation periods studied and the proxies used for the market excess return and the risk-free rate are discussed.

4.1.1 Data source

The source of the share data used in this study is Sharenet. The prices are back-adjusted for rights issues and stock splits and returns have been calculated ignoring dividends. The exclusion of dividends should have little impact. Evidence provided by Sharpe and Cooper (1972) shows that the value of beta does not change significantly if dividends are excluded. Their study found an almost perfect relationship (correlation coefficient of 0.997) between the two types of beta coefficients. Bloombergs was used as a check to ensure the correctness of both price levels and volumes traded. All shares that were listed continuously on the Johannesburg Stock Exchange for the period 1 January 1988 to 17 March 1999 were considered for inclusion in the study. A total of 346 such shares were identified, i.e. shares that were both listed over the required period and available on Sharenet. Out of this pool of shares, a random sample of 100 shares was selected.

4.1.2 Partitioning on the basis of liquidity

The selected shares were classified into three groups: Liquid (L), illiquid (I) and an in-between group (B). This was done because earlier empirical studies indicated that the performance of estimators depends on liquidity considerations (see e.g. Bartholdy and Riding 1994, Dimson and Marsh 1983 and Fowler, Rorke and Jog 1980a&b). The frequency, with which the price of a share changes, was used as a proxy for liquidity in this case. For a share to be classified as liquid, its price had to change on more than one third (33%) of all the trading days in the period under review (1 January 1988 to 17 March 1999) AND its price had to change in more than 65% of the weeks in the period. Illiquid shares were defined as those whose prices changed in less than a quarter (25%) of the days during the whole period under review and in less than 55% of the weeks in the period. The remaining shares were classified as belonging to the in-between group. The result of this method of classification is that 35 shares were classified as liquid, 31 as illiquid, and 34 to the in-between group. The trading frequency of a share could also be used as a measure of liquidity. It could be argued, however, that it is a characteristic of illiquid shares that they sometimes trade very thinly, but that their prices normally do not change when traded. Liquid shares, on the other hand, normally have high volumes on any day and it is seldom that their prices do not change on this volume. If active days were used as a measure of liquidity, high and low volumes would have equal weights.

4.1.3 Return calculation periods

Three return calculation periods were used. The reason for studying different calculation periods was partly motivated by similar empirical studies conducted in the past. These studies (see e.g. Chan and Lakonishok 1992 and Draper and Paudyal 1995) indicated that estimators behave differently when returns are calculated over different time periods. Daily, weekly, and monthly returns were calculated in the following way:

$$R_{it} = \left(\frac{P_{i,t}}{P_{i,t-1}}\right) - 1 \qquad \dots (8)$$

where $P_{i,t}$ is the price of share i at time t, and $P_{i,t-1}$ is the price of the share one period earlier. Weekly returns were calculated from Friday to Friday (or the first business day after the Friday where a particular Friday was not a business day). Monthly returns were calculated from the last day of a month to the last day of the next month (or first business day after the last day of a month where such a day was not a business day).

4.1.4 The market proxy and the proxy for the risk-free rate

We used the JSE/Actuaries All Share Index as proxy for the market. The objective of the JSE/Actuaries All Share Index (All Share) is to include all securities exclusive of pyramid companies, debentures and preference shares.

As mentioned before, the single index model (where excess stock returns are regressed against excess market returns) is considered. That is, the risk-free rate of return is subtracted from the return on individual stocks and the proxy for the market index. There are several possible interest rates that can be used as a proxy for the short-term risk-free rate in the South African case. The more important ones are the BA rate, the 3-month NCD rate, the prime overdraft rate, the overnight interbank lending rate, the repo rate and the 3 month TB rate. The rate used in this study is the BA rate. Over the period under review the BA rate has constantly been the benchmark short-term rate in South Africa. The data used are from Ecoserve, which is an in-house database used and maintained by Standard Bank. Bloombergs was again used to check data for correctness and validity. Note that the annual interest rates had to be adjusted to the applicable period (daily, weekly, monthly). This was done (in the case of the daily data set) by dividing each day's value by 365. For weekly data the value on the particular day was divided by 52 and for monthly data by 12.

4.2 The estimators and the performance measures

4.2.1 The estimators

Apart from OLS, three robust L-estimators are studied, viz. the L_1-estimator (L1), the Koenker-Bassett regression trimmed mean (KB), and the Welsh trimmed mean (WE). The bounded-influence versions of these estimators are also used, viz. the bounded-influence L_1-estimator (BI-L1), the bounded-influence Koenker-Bassett estimator (BI-KB), and the bounded-influence Welsh estimator (BI-WE). We will refer to the last six estimators as the "robust estimators". Note that all the trimmed mean estimators were calculated using a trimming proportion of 15%. Seven additional estimators can be formed by basing the Vasicek estimator on OLS and on the six robust estimators. These are denoted by VAS-OLS, VAS-L1, VAS-WE, VAS-KB, VAS-BI-L1, VAS-BI-WE, and VAS-BI-KB. As mentioned previously, partitioning was done on the basis of the liquidity of a share. The three classes of liquidity are the liquid, in-between and illiquid. Rather than computing the Vasicek estimators on all hundred shares, the calculation can also be based on those shares belonging to a specific liquidity class. The seven partitioned Vasicek estimators that can be formed in this way are denoted by VAS-OLS (class), VAS-L1 (class), VAS-WE (class), VAS-KB (class), VAS-BI-L1 (class), VAS-BI-WE (class), and VAS-BI-KB (class). In total 21 estimators are used in the study.

4.2.1 The performance measures

In order to evaluate the performance of the various estimators in terms of forecasting stability, the Theil mean squared forecast error (MSE) was used (see e.g. Klemkosky and Martin 1975a&b and Murray 1995). This mean squared forecast error is defined as follows:

$$MSE = \frac{1}{m}\sum_{i=1}^{m}(A_i - P_i)^2 \qquad \ldots (9)$$

where:
m is the number of securities for which beta is predicted,

P_i is the prediction of the beta coefficient of security i, and

A_i is the actual estimated beta coefficient of security i.

In terms of the beta forecast, P_i (later referred to as predictions) represents the computed beta for the current period used as the predictor of beta for the subsequent period and A_i (later referred to as realisations) is the corresponding estimated beta for the subsequent period. MSE was chosen as the measure of forecast error because of its intuitive appeal and general acceptance, but also because it can be easily partitioned into three components of forecast error. This helps to identify the sources of forecast errors of extrapolated beta coefficients. MSE in (9) can be broken down as follows:

$$MSE = \left(\overline{A} - \overline{P}\right)^2 + (1-\beta_1)^2 S_P^2 + \left(1 - r_{AP}^2\right) S_A^2 \qquad \ldots (10)$$

where:
$\overline{A}$ and $\overline{P}$ are the means of the realisations and predictions,

β_1 is the slope coefficient of the regression of A on P,

S_P^2 and S_A^2 are the sample variances in P and A, and

r_{AP}^2 is the coefficient of determination for P and A. The first term in equation (10) represents bias squared, the second term inefficiency and the third is the random disturbance component of MSE. Bias in a forecast indicates that the average prediction was either over or under the average realisation. Inefficiency in the forecast represents a tendency for the prediction errors to be positive at low values of P_i and negative at high values of P_i. Put differently, it represents a tendency of the beta values to drift towards their mean. The beta adjustment methods of Blume and Vasicek typically attempt to minimise inefficiency in forecasts. The remaining component of MSE is the random disturbance which contains those forecast errors not related to the value of the predictor, P_i, or the predicted, A_i. This term tests the accuracy of the betas as predictors, as it essentially depends on the magnitude of the regression coefficient, r_{AP}^2.

Apart from using Theil's MSE, we also investigated the performance of the estimators by calculating the interquartile ranges (IQR) and median absolute deviations (MAD) of the residuals from each fit. Estimators having smaller IQR and MAD on average can be judged more efficient as than those having larger IQR and MAD on average.

In order to gain an understanding of the variability between the various estimators, we also calculated the range of the beta estimates obtained when using the different estimators.

It should be noted that, because the true underlying beta is unknown, it is notoriously difficult to compare estimators using empirical data. The above-mentioned measures are frequently used in empirical studies and serve as a gauge against which one can study the performance.

4.2.3 The data sets

For daily return data, 11 periods of 1 year each were used, viz. 1 January 1988 to 31 December 1988, 1 January 1989 to 31 December 1989, and so on until 1 January 1998 to 31 December 1998. Note that about 250 observations are used to calculate beta. For weekly returns, 5 periods of 2 years each were used: 1988-1989, 1990-1991, 1992-1993, 1994-1995 and 1996-1997. For monthly returns, 3 periods of 3 years each were used: 1988-1990, 1991-1993 and 1994-1997. Beta coefficients computed in one period were used to predict beta for the subsequent period and the MSE was used as performance measure. The IQR and MAD were averaged over the periods considered, and reported for each estimator separately (per liquidity classification and overall). The ranges were averaged over the periods considered and reported similarly.

4.3 Results of the empirical study

Using Theil's MSE as measure of forecasting performance, the performance of the estimators are now studied. We only present the results for the daily and monthly returns, because the conclusions drawn from the weekly returns are similar to those for the daily returns.

4.3.1 Daily returns

In Figure 1 the results obtained for the daily returns are presented. For all 21 estimators the average MSE, divided into average squared bias, average inefficiency, and the average random disturbance component, are shown. The average is obtained by averaging the above-mentioned quantities across all the prediction periods applicable. As each period's beta is used as an estimate for the beta of the next period, there has to be one less prediction period than there are time periods available. In this case 10 prediction periods for the daily returns. We used the average MSE to facilitate easy interpretation of the results. Detailed results are available from the corresponding author.

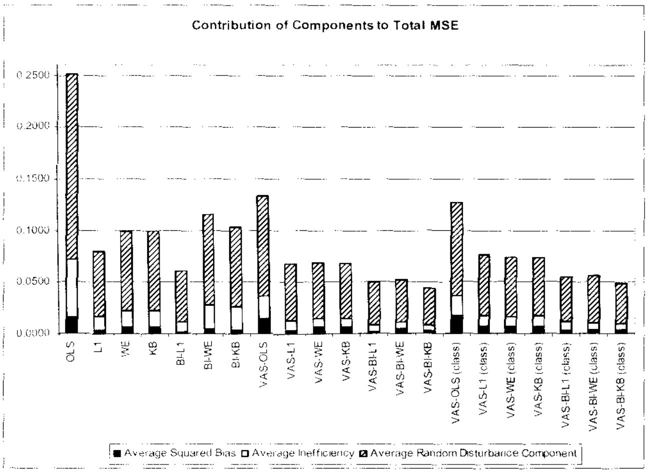

Figure 1: Average MSE's for the 21 estimators studied for the daily returns

In terms of average MSE, the OLS estimator is clearly the worst estimator of the 21 estimators considered. Also, the Vasicek-estimator based on OLS does not perform well when compared to the robust estimators and to the Vasicek-estimator based on the robust estimators. The partitioned Vasicek-estimators are generally performing on an equal footing with the unpartitioned Vasicek-estimators. The best estimators seem to be the VAS-BI-KB, VAS-BI-L1, VAS-BI-KB(class), VAS-BI-WE, VAS-BI-L1(class) and VAS-BI-WE(class). Thus, Vasicek and partitioned Vasicek estimators based on bounded-influence estimators are the best performers. The good performance of bounded-influence estimators can be attributed to the presence of extreme market returns. Amongst the robust estimators, BI-L1 and L1 performed best. It is important to note that when we consider each of the ten prediction periods, the Vasicek-estimator based on robust estimators, outperformed the Vasicek-estimator based on OLS over all the periods considered.

4.3.2 Monthly data

In Figure 2 the results obtained for the monthly returns are presented. For all 21 estimators the average MSE, divided into average squared bias, average inefficiency, and the average random disturbance component, are shown. The average is obtained by averaging the above-mentioned quantities across all the prediction periods applicable. As each period's beta is used as an estimate for the beta of the next period, there has to be one less prediction period than there are time periods available.

For monthly data the dominance of the robust methods over OLS is not as clear-cut as is the case for daily and weekly data. OLS even outperforms some of the robust estimators in terms of total MSE (e.g. BI-L1 and BI-WE). Again there seems to be no real difference between the Vasicek and partitioned Vasicek estimators. The best performers were VAS-KB(class), VAS-OLS, VAS-KB, VAS-WE(class), VAS-WE and VAS-OLS(class). Thus, Vasicek and partitioned Vasicek estimators performed well, but not those based on bounded-influence estimators. This is contrary to what was found for daily and weekly data. This can be attributed to the fact that monthly return data smooth out the presence of extreme returns. Four estimators clearly performed worse than the others: BI-L1, BI-WE, OLS and BI-KB. An interesting observation is that VAS-BI-L1 and VAS-BI-L1(class) performed badly relative to the other Vasicek estimators that were used.

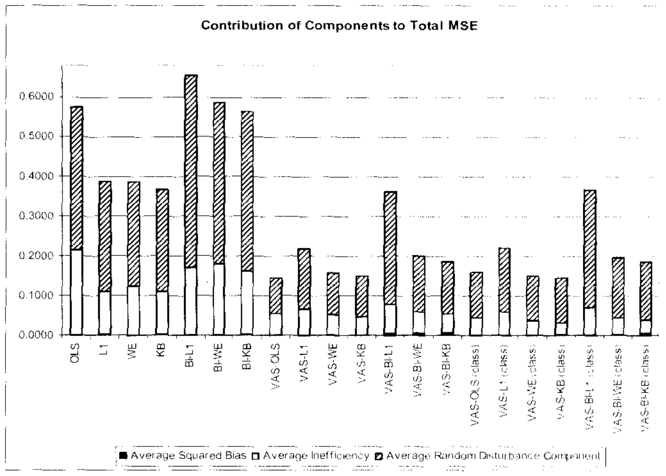

Figure 2: Average MSE's for the 21 estimators studied for the monthly returns

Overall the results showed that for monthly returns that OLS, and especially the bounded-influence estimators was by far the worst estimators amongst those examined and that significant benefits can be achieved by using Vasicek's estimator based on OLS and robust estimators, but not on bounded-influence estimators.

In Table 1 below we present a summary table representing the drop in average total MSE obtained relative to OLS. Because the results of the WE and KB trimmed means are very similar we only show the results obtained for the KB trimmed means.

Table 1: Average total MSE of estimators expressed as a percentage relative to the average total MSE of OLS

Estimators	Daily	Weekly	Monthly
OLS	100	100	100
KB	40	35	64
BI-KB	41	37	98
VAS-OLS	53	33	25
VAS-KB	28	23	26
VAS-BI-KB	18	21	33

From Table 1, the percentage drop in average total MSE obtained by using KB instead of OLS for daily data is 60% (OLS is 100% and KB is 40% of OLS). Clearly, for daily data, the use of robust estimators leads to a substantial gain in prediction accuracy. Similarly VAS-BI-KB for daily data resulted in an 82% drop (100% to 18%) in average total MSE compared to OLS. Relative to the best robust estimator (KB), this result in a further 55% ((40-18)/40) improvement in average total MSE. Again, it is clear that for daily returns, the new estimators do very well, especially the Vasicek-estimator based on the bounded-influence trimmed mean (VAS-BI-KB). For weekly data the results are similar, except that the performance of the Vasicek-estimator based on OLS is better than what was the case for daily return data. For weekly data it now outperforms the robust estimators, which was not the case for daily data. However, for both the daily and weekly data, the new estimators are still the best performers. For monthly data, the new estimators do well, but is slightly outperformed by the Vasicek-estimator based on OLS. Note the poor performance of the robust and bounded-influence estimators.

4.3.3. The effect of liquidity on MSE

It was found that the Vasicek-estimators based on the bounded-influence estimators performed best in terms of average MSE over the period considered. For the daily return data of illiquid shares, the Vasicek-estimators based on the robust and bounded-influence estimators perform much better than the Vasicek-estimator based on OLS. In the case of illiquid shares the improvement, in relative terms, was found to be much better than in the case of liquid shares. This is the result of the fact that the prices of most of the shares in the illiquid class do not change over most of the periods considered. This conclusion also holds true for the weekly and monthly data, where the partitioned Vasicek-estimators based on robust and bounded-influence estimators performed best for the illiquid class of shares.

4.3.4 MAD, IQR and range

The average MAD and IQR support the above-mentioned findings. In general the Vasicek estimator based on robust estimators outperform the other estimators for daily return data. For monthly data, the performance of all the estimators is very similar. As far as the variation between beta estimates is concerned, the ranges calculated indicate that the Vasicek estimators based on the robust estimators provide similar estimates, i.e. the average ranges calculated are small in comparison to the average ranges calculated when the other estimators are included. This gain in stability is much more marked for daily return data than for monthly return data.

5. CONCLUDING REMARKS

The results of the study clearly show that the idea of combining robust estimators with the Vasicek-estimator yields a class of new estimators that performs well when compared to traditional estimators. The improvement in performance achieved is very clear for daily returns and especially so in the illiquid category. For all the other factors (weekly and monthly return data and the other liquidity categories) considered the estimators performed well, but their superior performance is not as prominent. Especially in the case of the monthly return data and the liquid category, the new estimators do not outperform the traditional Vasicek-estimator.

A natural next step would be to use the above-mentioned estimators to estimate and forecast the covariance structure between the shares (see e.g. the study by Elton, Gruber and Urich, 1978). This is an important input to mean variance analysis and the performance of the estimators can be studied in terms of the accuracy with which the covariance structure is forecast and how accurate the optimal portfolio is determined. One could also study the performance of the estimators by means of a simulation study. In such a study the behaviour of the returns of the individual shares could be modelled in each liquidity category and then the performance of the various estimators can be evaluated in terms of the above-mentioned performance measures.

REFERENCES

Bartholdy J and Riding A. (1994). Thin trading and the estimation of betas: The efficacy of alternative techniques. *Journal of Financial Research*, 17(2):241-253.

Blume ME. (1973). Beta's and their regression tendencies. *Journal of Finance*, 28:785-795.

Boabang F. (1996). An adjustment procedure for predicting betas when thin trading is present: Canadian evidence. *Journal of Business Finance and Accounting*, 23(9/10):1333-1355.

Bowie DC and Bradfield DJ. (1998). Robust estimation of beta coefficients: Evidence from a small stock market. *Journal of Business Finance and Accounting*, 25(3/4):439-454.

Chan LKC and Lakonishok J. (1992). Robust measurement of beta risk. *Journal of Financial and Quantitative Analysis*, 27:265-282.

De Jongh PJ. (1985). The regression trimmed mean: An empirical study. Ph.D-thesis, University of Cape Town.

De Jongh PJ and de Wet T. (1986). Confidence intervals for regression parameters based on trimmed means. *South African Statistical Journal,* 20(2):137-164.

De Jongh PJ, de Wet T and Welsh AH. (1988). Mallows-type bounded influence regression trimmed means, *Journal of the American Statistical Association*, 83(403):805-810.

Dimson E and Marsh P. (1983). The stability of UK risk measures and the problem of thin trading, *Journal of Finance*, 38:753-783.

Draper P and Paudyal K. (1995). Empirical irregularities in the estimation of beta: The impact of alternative estimation assumptions and procedures. *Journal of Business Finance and Accounting,* 22(1):157-177.

Elton EJ and Gruber MJ. (1995). Modern portfolio theory and investment analysis, 5th ed. New York, John Wiley and Sons, 715 p.

Elton EJ and Gruber MJ. (1997). Modern portfolio theory, 1950 to Date. *Journal of Banking and Finance*, 21:1743-1759.

Elton EJ, Gruber MJ and Urich TJ. (1978). Are betas best?', *Journal of Finance*, 33:1375-1384.

Fowler D, Rorke CH and Jog VM. (1980a). A bias-correcting procedure for beta estimation in the presence of thin Trading. *Journal of Financial Research*, 12:23-32.

Fowler D, Rorke CH and Jog VM. (1980b). 'Thin trading and beta estimation problems on the Toronto Stock Exchange. *Journal of Business Administration,* (Fall 1980):77-90.

Klemkosky RC and Martin JD. (1975a). The adjustment of beta forecasts. *Journal of Finance*, 30(4):1123-1128.

Klemkosky RC and Martin JD. (1975b). The effect of market risk on portfolio diversification. *Journal of Finance*, 30(1):147-154.

Koenker RW and Bassett G. (1978). Regression Quantiles. *Econometrica*, 46:33-50.

Koenker RW and Machado JAF. (1999). Goodness-of-fit and related inference processes for quantile regression. *Journal of the American Statistical Association*, 94(448):1296-1307.

Lam KSK. (1999). Some evidence on the distribution of beta in Hong Kong. *Applied Financial Economics*, 10:251-262.

Lally M. (1998). An examination of Blume and Vasicek betas. *Financial Review*, 33:183-198.

Murray L. (1995). An examination of beta estimation using daily Irish data. *Journal of Business Finance and Accounting*, 22(6):893-905.

Ruppert D and Carroll RJ. (1980). Trimmed least squares estimation in the linear model. *Journal of the American Statistical Association*, 75:828-838.

Scholes M and Williams J. (1977). Estimating betas from non-synchronous data. *Journal of Financial Economics*, 5:309-328.

Sharpe W and Cooper G. (1972). Risk-return classes of New York Stock Exchange common stocks. *Financial Analysts Journal*, 28(2):46-54.

Vasicek O. (1973). A note on using cross-sectional information in Bayesian estimation on security beta's'. *Journal of Finance*, 28(5):1233-1239.

Welsh AH. (1987). The trimmed mean in the linear model. *Annals of Statistics*, 15:20-45.