Estimates of Small-Stock Betas are Often Very Distorted by Outliers

Robust betas are not much influenced by outliers and provide a viable complement to the OLS beta.

R. Douglas Martin and Tim Simin

TECHNICAL REPORT NO. 351

May 1999

Department of Statistics
Box 354322
University of Washington
Seattle, WA 98195 USA

Estimates of Small-Stock Betas are Often Very Distorted by Outliers

Robust betas are not much influenced by outliers and provide a viable complement to the OLS beta.

R. Douglas Martin and Tim Simin*

ABSTRACT

The mounting evidence of outlier generating heavy-tailed or mixture distribution functions for equity returns motivates the search for estimation procedures robust toward outlying data points. This paper introduces such an estimator and presents a detailed comparison with the classical ordinary least squares (OLS) estimator using historical monthly equity returns data from the CRSP database. The results show that the two estimates differ significantly for a non-trivial fraction of the equities studied. Furthermore, such behavior occurs primarily for firms with small market capitalization. The results of our study suggest that robust estimates of beta may be of considerable value as a complement to the standard OLS beta estimates.

R. DOUGLAS MARTIN is a Professor of Statistics at the University of Washington and Chief Scientist of the Data Analysis Products Division of MathSoft, Inc, both in Seattle, WA. (Email: doug@statsci.com).

TIM SIMIN is a Ph.D. candidate in the Department of Finance and Business Economics, School of Business, University of Washington, Seattle, WA. (Email: tsimin@u.washington.edu).

1. INTRODUCTION

The estimated slope coefficient of $\boldsymbol{\beta}$ in the classic market model

$$R_t = \boldsymbol{\alpha} + R_{m,t} \cdot \boldsymbol{\beta} + \boldsymbol{\varepsilon}_t , \quad t = 1, \cdots, T \tag{1}$$

is undoubtedly the best known and most frequently used measure of risk and return. Here, R_t is the time series of a particular equity's returns and $R_{m,t}$ is the time series of returns for a market proxy, both in excess of the risk-free interest rate for monthly data, or perhaps as raw returns for weekly data.

Virtually all published sources of beta, and indeed the vast majority of academic studies, use either the raw or an adjusted version of the ordinary least squares (OLS) estimate of $\boldsymbol{\beta}$. The sanctified use of OLS for statistical modeling and inference is justified by the fact that OLS is the best linear unbiased (BLUE) estimate of linear model coefficients, and the overall (linear or nonlinear) best estimate when the errors are Gaussian. Furthermore, if the errors are Gaussian OLS furnishes a convenient distribution theory for inference.

However, since the seminal work by Mandelbrot (1963) empirical investigations have produced evidence that equity returns contain outliers generated by heavy-tailed probability distributions. The early work of Fama (1965) provided evidence in favor of stable distributions. Later on, authors such as Kon (1984), Roll (1988), Connolly (1989), and Richardson and Smith (1994) provided evidence in favor of the use of normal mixture distributions for generating outliers in financial returns.[1]

It is well known in the statistics literature that outliers generated by heavy-tailed distributions often have a substantial distorting influence on least squares estimates. In such situations, the OLS estimate is no longer the best estimates in the class of all linear and nonlinear estimates. In particular, outlier-generating distributions can cause the OLS estimate to suffer from large biases in the presence of outliers.[2] It is therefore important to have alternative *robust* estimators that are not much influenced by outliers and still perform nearly as well as OLS when the data is outlier free and has a normal distribution. This issue has been extensively studied in the statistical literature, with regard to both theory and applications.[3]

In view of the above, it is perhaps surprising that the use of robust estimators has been given relatively little attention in the finance literature overall, and in the context of estimating beta in particular. This may be partly due to the negative results for robust beta estimation initially (Sharpe, 1971), and by Cornell and Dietrich (1978), and the somewhat unimpressive positive results of several subsequent studies.[4]

However, Knez and Ready (1997) recently used robust regression to establish the striking result that the negative risk premium on size, reported by Fama and French (1992) using the OLS estimate, is caused by a small fraction of outliers (less than 1%) in the data. These outliers typically consist of exceptionally large returns for small sized firms.

Furthermore, eliminating these few outlying firm returns produces a positive relationship between average returns and firm size. Knez and Ready's use of robust regression was most impressive in providing a clear, succinct and accurate model for describing the relationship between equity returns and firm size, namely: For most firms, the relationship between returns and firm size is positive, but there exist a very small fraction of small-sized firms that have exceptionally large outlier returns.

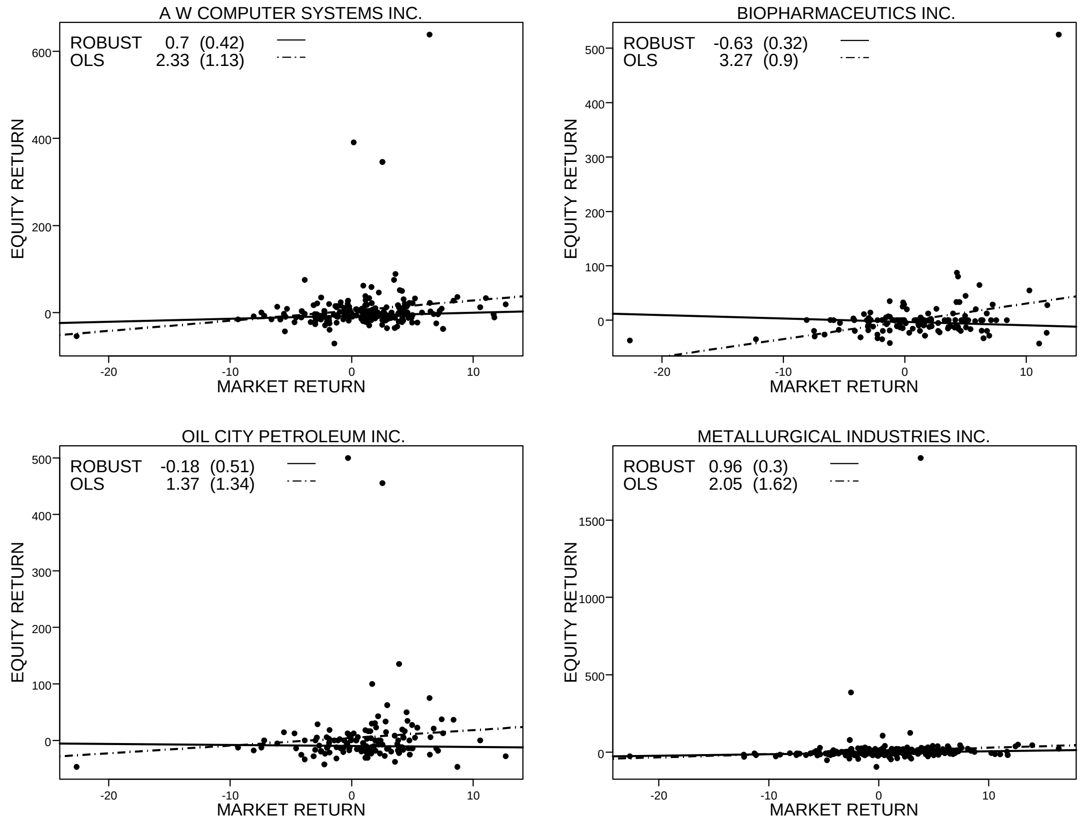
Motivated by the Knez and Ready's highly successful use of a robust regression estimate, we propose using a new robust beta estimate that is well suited for dealing with outliers that distort the OLS estimate of beta in the market model (1). We carry out an extensive study of the comparative behavior of the OLS and robust betas on monthly equity returns from the Center for Research on Stock Prices (CRSP) database. The results reveal that the difference between the OLS and robust beta estimates sometimes differ by more than one in magnitude, and frequently by more than one-half. Differences of this magnitude are likely to be financially significant in many contexts. Our results also reveal that substantial outlier distortion of the OLS beta occurs primarily for small-sized firms. This small-firm effect is consistent with the findings of Knez and Ready (1997) study.

2. SOME EXAMPLES COMPARING OLS AND ROBUST BETA'S

Before describing our proposed robust estimate, we provide four (of many) striking examples, where there is a large difference between the OLS and robust beta due to outlier-induced distortion of the OLS beta.

Exhibit 1 displays the scatter plots of monthly equity excess returns versus excess returns for a market proxy for four small firms. The OLS beta estimate and our proposed robust beta estimates are the slopes of the dashed and solid straight line fits respectively. The outliers have considerable influence on the OLS beta causing it to be a relatively poor fit to the bulk of the data. The robust beta is not much influenced by the outliers and provides a good fit to the bulk of the data, which is an important feature of a good robust estimate. The market proxy used in this exhibit, and throughout the paper, is the NYSE, AMEX, and NASDAQ composite, with returns being in excess of the one-month T-bill rate from the CRSP database. The returns are for the time interval from January 1962 to December 1996.

EXHIBIT 1
OLS AND ROBUST ESTIMATES OF BETA FOR FOUR SMALL FIRMS



The OLS and robust beta estimates of Exhibit 1 are assembled in Panel A of Exhibit 2. In each of the four examples the OLS beta is greater than one, substantially so in three out of the four cases, and the values are all considerably larger than the robust beta by almost any standard of comparison. This positive *bias* in the OLS estimate is caused by a general tendency for large positive outliers to occur when the market returns are more or less positive.

EXHIBIT 2
OLS AND ROBUST BETA ESTIMATES

Panel A: Beta Estimates, (std. errors)				
	A W Computer Systems Inc.	Oil City Petroleum Inc.	Biopharmaceutics Inc.	Metallurgical Industries Inc.
OLS	2.33	1.37	3.27	2.05
	(1.13)	(1.34)	(0.90)	(1.62)
ROBUST	0.70	-0.18	-0.63	0.96
	(0.42)	(0.51)	(0.32)	(0.30)
Panel B: Beta Estimates with Oct. 1987 Deleted, (std. errors)				
	A W Computer Systems Inc.	Oil City Petroleum Inc.	Biopharmaceutics Inc.	Metallurgical Industries Inc.
OLS	2.27	1.1	3.67	2.11
	(1.24)	(1.55)	(1.01)	(1.71)
ROBUST	0.70	-0.18	0.68	0.95
	(0.43)	(0.05)	(0.47)	(0.33)

In all four examples the crash of October 1987 (the data point in the lower left corner of each graph) is a *leverage* point, i.e., an outlier in the excess market returns[5]. One is naturally concerned that this leverage point outlier might substantially contribute to increasing the value of the OLS beta over that of the robust beta. However, Panel B of Exhibit 2, shows quite clearly that the OLS estimate is relatively unaffected by the October 1987 outlier. It is the other outliers that are causing the OLS beta to be substantially larger than the robust beta.

We emphasize that the behavior exhibited by the four examples of Exhibit 1 and Exhibit 2 typifies a general tendency observed in other cases where there is a large difference between the OLS and robust beta. Namely, the outliers often fall in configurations that cause the OLS beta to be biased toward larger values than those of the robust beta. In the presence of outliers, the OLS estimate of beta is *too high* more often than not. This tendency is particularly strong for the monthly returns results reported in a later section.

It is important to be aware that outliers can substantially influence not only the OLS estimates of the linear model parameters, but also the estimated standard errors of these parameter estimates. Because the standard error estimate uses an error variance estimate based on the sum-of-squared residuals, residual outliers inflate the standard error estimate. On the other hand, the robust estimate down-weights outliers as described subsequently, and hence the robust estimate yields an estimate of standard errors which is insensitive to outliers. The above facts are clearly illustrated in Exhibit 2 where the robust beta standard error estimates are much smaller than the OLS beta standard error estimates.[6]

Interpretation of OLS and Robust Beta

The presence of outlier induced positive bias in the OLS betas mistakenly portrays three of the four firms in Exhibit 1 the as aggressive assets in the sense that they have a high level of risk and expected return. This does not sufficiently characterize true nature of these firms. First of all, the behavior of the overwhelming bulk of the data suggests levels of risk and return at or below that of the market. Secondly, the returns giving rise to the unusually large values of OLS betas represent a very small number of unusually large positive returns that tend to occur when the market returns are more or less positive, with no corresponding negative returns in evidence. Thus, on the basis of the returns data alone it is hardly a complete summary to simply state, as suggested by the OLS beta, that the firms in question are more risky than the market. A more accurate description would be to say that for the most part, the firms risk and return characteristics are represented by the robust beta, and that in addition there are a very small fraction of unusually large positive returns with no unusually negative returns in evidence. Such firms may be particularly interesting to investigate further with regard to investment opportunities.

A fundamental point suggested by the above examples, and many others like it in our study, is that neither an OLS nor a robust beta estimate alone suffices to categorize the risk and return characteristics of a firm. When the two estimates agree it signals the absence of distortion of the OLS beta by influential outliers, and in such cases one can reasonably trust the classical interpretation of the OLS beta. When the two estimates differ significantly according to either a user specified threshold or an appropriate test statistic, it signals the presence of influential outliers.[7] In such cases the robust beta characterizes the behavior of the bulk of the data, but additional investigation is warranted. For example, the residuals from the robust linear fit could be used to automatically identify the sizes and occurrence times of outliers, and the occurrence times and time patterns of outliers may help the investor identify their causes.

3. THE PROPOSED ROBUST BETA

Our proposed robust beta is obtained using a regression M-estimate approach that is quite well-known in the statistical literature.[8] M-estimates $\hat{\alpha}$ and $\hat{\beta}$ of the intercept and slope coefficients α and β in (1) are obtained by minimizing the quantity

$$\sum_{t=1}^{T} \rho\left(\frac{R_t - \alpha - R_{m,t} \cdot \beta}{\hat{s}}\right) \qquad (2)$$

with respect to α and β. Here ρ is a symmetric *loss* function, and $\hat{s}$ is a robust scale estimate for the residuals. The latter makes the estimates of intercept and slope invariant with respect to the scale of the errors ε_t in (1). The OLS and LAD estimates are special cases of (2) corresponding to quadratic and absolute value loss functions, respectively.[9]

The estimating equations, obtained by differentiating (2) with respect to $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$ and setting $\psi = \rho'$, are

$$\sum_{t=1}^{T} \boldsymbol{\psi}\left(\frac{R_t - \hat{\boldsymbol{\alpha}} - R_{m,t} \cdot \hat{\boldsymbol{\beta}}}{\hat{s}}\right) = 0 \tag{3}$$

and

$$\sum_{t=1}^{T} R_t \cdot \boldsymbol{\psi}\left(\frac{R_t - \hat{\boldsymbol{\alpha}} - R_{m,t} \cdot \hat{\boldsymbol{\beta}}}{\hat{s}}\right) = 0\,. \tag{4}$$

These estimating equations have a simple weighted-least-squares (WLS) interpretation as follows. Define residuals

$$\hat{\boldsymbol{\varepsilon}}_t = R_t - \hat{\boldsymbol{\alpha}} - R_{m,\mathrm{t}} \cdot \hat{\boldsymbol{\beta}} \tag{5}$$

and *data-dependent* weights

$$w_t = \frac{\boldsymbol{\psi}\left(\frac{\hat{\boldsymbol{\varepsilon}}_t}{\hat{s}}\right)}{\hat{\boldsymbol{\varepsilon}}_t} \cdot \hat{s}\ , \quad t = 1, \cdots, T\ . \tag{6}$$

These weights are the result of applying the weight function

$$w(u) = \frac{\boldsymbol{\psi}(u)}{u} \tag{7}$$

to the standardized residuals $\dfrac{\hat{\boldsymbol{\varepsilon}}_t}{\hat{s}}$.

Now the estimating equations (3) and (4) may be written in the WLS form

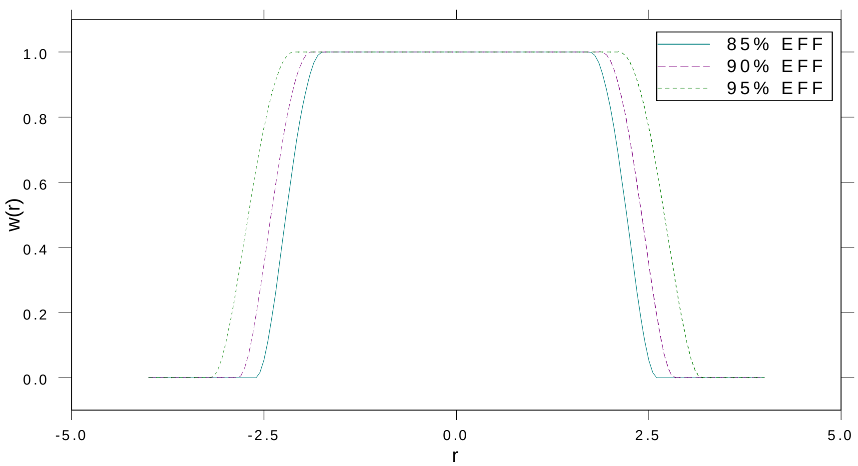
$$\sum_{t=1}^{T} w_t \cdot \left(R_t - \hat{\boldsymbol{\alpha}} - R_{m,t} \cdot \hat{\boldsymbol{\beta}}\right) = 0 \tag{8}$$

$$\sum_{t=1}^{T} R_t \cdot w_t \cdot \left(R_t - \hat{\boldsymbol{\alpha}} - R_{m,\,t} \cdot \hat{\boldsymbol{\beta}}\right) = 0\,. \tag{9}$$

For our robust beta estimates we use a weight function $\boldsymbol{w}$ obtained from a particular type of *bounded* loss function $\boldsymbol{\rho}$, and corresponding $\boldsymbol{\psi} = \boldsymbol{\rho}'$, that provides good protection against bias due to outliers and has a user-specified efficiency when the errors $\boldsymbol{\varepsilon}_t$ in the market model (1) are Gaussian.[10] Yohai and Zamar (1997) recently established the robustness properties of this loss function.

Several weight functions obtained from the Yohai and Zamar type loss function are shown in Exhibit 3. The three cutoff values of c = 0.868, 0.944, and 1.060, at which the weight function goes to zero, yield efficiencies of 0.85, 0.90, and 0.95 when the errors in (1) are Gaussian. Increasing values of c give increased efficiency when the data is Gaussian, but less protection against bias from outliers that are not generated by the market model (1). The formula for the weight function $\boldsymbol{w}$ is given in the Appendix.

EXHIBIT 3
ROBUST WEIGHT FUNCTIONS FOR SEVERAL GAUSSIAN EFFICIENCIES



The weights in Exhibit 3 have the following intuitive interpretation. They are smooth approximations to the so-called *hard-rejection w*-function, which has a constant value of unity on the interval (-c, c) and has value zero outside that interval. The smooth transition of w from the full-weight value of one in the central region to zero outside the interval (-c, c) is intuitively appealing and preferable to the discontinuous behavior of the hard-rejection weight function. The effect of the latter is to regard all sufficiently small observations to be "completely" good, and all others to be "completely" bad, which is patently unnatural. On the other hand, the region of smooth transition of w in Exhibit 3 occurs where most natural, namely in the "flanks" of the distribution where it is most difficult to decide whether or not an observation is an outlier.

Throughout this paper we use the value c = 0.944 which yields a 90% variance efficiency for normal errors. This corresponds to 94.9% efficiency on a standard deviation basis, i.e., the standard deviation of the robust parameter estimates is about 5.4% greater than that of OLS when we use c = .944. This yields good protection against bias due to outliers that do not conform to the market model (1), at a very small premium in loss of efficiency when returns are Gaussian.

Computing the Robust Beta

The robust beta is obtained by minimizing (2) with a bounded and hence non-convex loss function ρ. Thus, computing the robust beta requires minimizing a function that may have multiple local minima. Equivalently, the robust beta is obtained by solving (8) and (9) with weights derived from a weight function that is zero outside ($-c$, c), and therefore (8) and (9) may have multiple roots. Fortunately, a reliable computational strategy for solving regression M-estimate minimization problems like (2) was provided by Yohai, Stahel and Zamar (1991). Their method has been implemented in the commercial data analysis and statistical modeling system S-PLUS (1998). We have used the S-PLUS implementation to compute robust betas in this paper.

Standard Errors of the Robust Betas

The lack of inference for robust regression parameter estimates has undoubtedly hindered the widespread use of robust estimators. A convenient feature of the robust M-estimate approach is that the parameter estimates are approximately normally distributed, with a covariance matrix that can be conveniently estimated yielding approximate standard errors, p-values and the like. We report the standard errors produced by the S-PLUS implementation in this paper. A detailed description of the computation of standard errors for the robust beta may be found in Martin and Simin (1999).

4. THE CRSP DATA SET COMPARISONS

We compare the behavior of the OLS and robust beta estimates using monthly returns for firms listed on the NYSE, AMEX, and NASDAQ exchanges from the CRSP database. We include firms that have been listed for at least 10-years of data during the period 1962 to 1996, and the series for each such firm extends over its entire listing period.

Exhibit 4 displays binned counts of the absolute and raw differences between the OLS and corresponding robust beta. The count for each bin is the number of firms for which the difference between the OLS and robust beta estimates fall in that cell. Most of the betas for the robust estimator are fairly close to the OLS beta. But significant fractions of the firms have OLS and robust betas that differ by moderate to large amounts. For example, 15.3% of the firms have differences larger than 0.3 and 5% have differences larger than 0.5. These are differences that are likely to be financially significant to many investors. It can be shown that a very large percentage of the differences, namely 23.9%, are statistically significant at a p-value of .01 (Martin and Simin, 1999).

EXHIBIT 4
BINNED COUNTS OF DIFFERENCES BETWEEN OLS AND ROBUST BETA'S

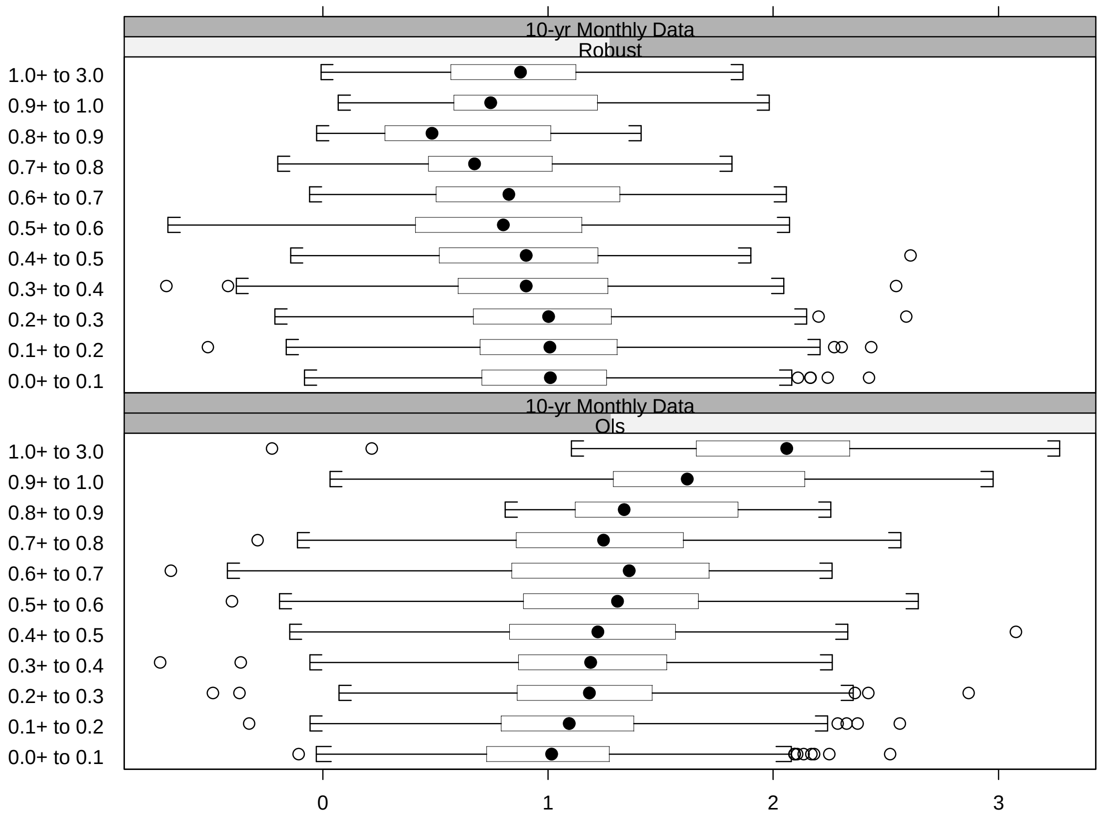
Panel A: $\Delta = \lvert\beta_{OLS} - \beta_{ROB}\rvert$		Panel B: $\Delta = \beta_{OLS} - \beta_{ROB}$	
0.0+ to 0.1	2909	< -0.8	3
0.1+ to 0.2	1493	-0.8+ to -0.6	19
0.2+ to 0.3	743	-0.6+ to -0.4	33
0.3+ to 0.4	410	-0.4+ to -0.2	151
0.4+ to 0.5	222	-0.2+ to 0.0	1403
0.5+ to 0.6	138	0.0+ to 0.2	2999
0.6+ to 0.7	59	0.2+ to 0.4	1002
0.7+ to 0.8	33	0.4+ to 0.6	327
0.8+ to 0.9	26	0.6+ to 0.8	73
0.9+ to 1.0	16	0.8+ to 1.0	41
> 1.0+	28	> 1.0	26
N =	6077	N =	6077

Furthermore, Panel B of Exhibit 4 clearly shows that the distorting influence of outliers on the OLS betas results in a positive bias, i.e., for symmetrically located pairs of bins, the positive bin counts are larger than the negative bins counts.

Exhibit 5 shows *Trellis* displays of conditional box plots of the OLS and robust beta estimates.[11] The vertical axis labels indicate the interval associated with the difference between the estimates, and the strip labels for each panel in the Trellis display indicate which estimate the box plots are based on.[12] For example, the strip labels indicate that the bottom panel is for the OLS beta and the top panel is for the robust beta. The upper strips are used to remind us that we are dealing with monthly data for firms that existed for at least 10 years in the CRSP database.

The Trellis display in Exhibit 5 immediately reveals the following behavior. For firms where the difference between robust and OLS betas is small, the distributions of the estimates look quite similar. This is consistent with the fact that for Gaussian data the OLS estimate is efficient, and the robust estimator has high efficiency and is therefore highly correlated with the OLS estimate. When the difference between the OLS and robust betas increases, the OLS beta distributions are shifted considerably in the positive direction and the robust beta distributions are shifted somewhat in the negative direction. This indicates that the larger biases in the OLS beta tend to occur for firms for which a few outliers cause a positive bias and for which the data has somewhat smaller beta values. This behavior deserves further investigation and explanation.

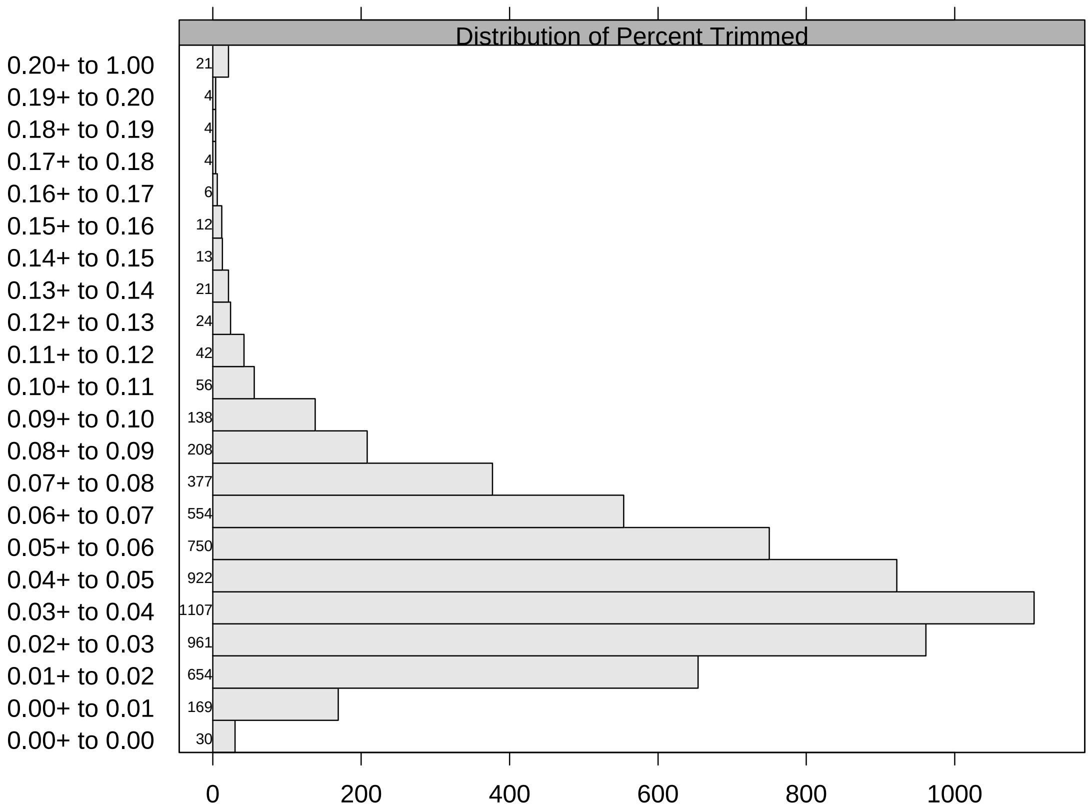
EXHIBIT 5
BOXPLOT VIEWS OF THE CONDITIONAL DISTRIBUTION OF OLS AND ROBUST BETA'S, CONDITIONED ON THE ABSOLUTE DIFFERENCE BIN



The Percentage of Outliers Rejected by the Robust Betas

One may reasonably ask how many data points are rejected by the robust estimator, where a data point at time *t* is considered rejected if its weight w_t in the weighted-least-squares estimating equations (8) and (9) is zero. Exhibit 6 below shows the distribution of the percentage of points rejected for the data used in our study.

EXHIBIT 6
DISTRIBUTION OF OUTLIERS REJECTED BY THE ROBUST BETA



The above exhibit shows that the most probable percentage of outlier rejections is between 3% and 4%, and that more than 10% of outliers are rejected only for the very small fraction 3.4% of the firms.

5. OUTLIERS DISTORTION OF OLS BETA IS A SMALL-FIRM EFFECT

Using the absolute difference between the OLS and robust beta's as an indicator of outliers influence on the OLS beta in a given firm data set, we examine the relationship between outlier influence on OLS and the average market capitalization (price times shares outstanding). Exhibit 7 displays location and spread statistics on market capitalization for each absolute difference bin. For each firm, the time series average over the sample period of the year-end price/share times the number of shares is used as the firm's market capitalization rate. Both the price per share and the shares outstanding are the values closest to the end of the year. The summary statistics are for the distribution of these market capitalization values on a bin-by bin basis.

EXHIBIT 7
MARKET CAPITALIZATION STATISTICS VERSUS ABSOLUTE DIFFERENCE BINS FOR OLS AND ROBUST BETAS*

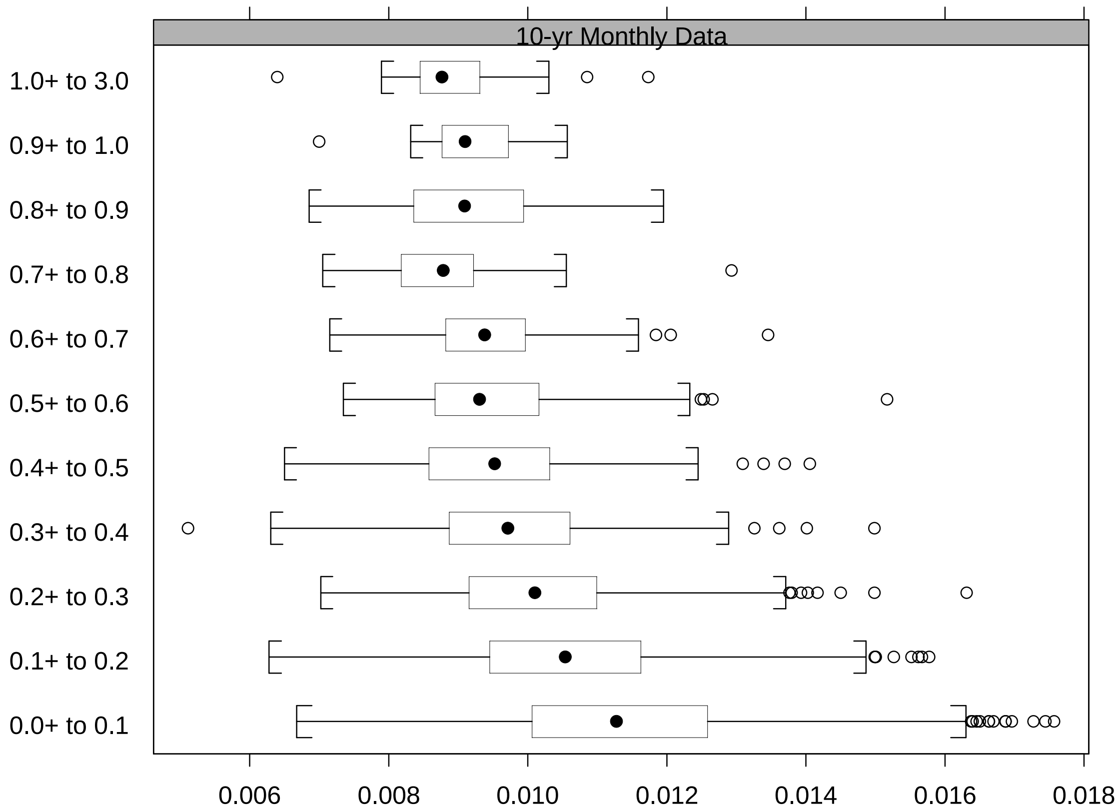
Bin	N	Mean	Std. Dev.
0.0+ to 0.1	2909	6894.48	23698.67
0.1+ to 0.2	1493	2342.87	8692.43
0.2+ to 0.3	743	1728.51	10685.68
0.3+ to 0.4	410	798.15	3775.63
0.4+ to 0.5	222	703.86	3388.03
0.5+ to 0.6	138	777.83	3554.71
0.6+ to 0.7	59	507.36	1158.51
0.7+ to 0.8	33	297.62	883.97
0.8+ to 0.9	26	284.14	409.99
0.9+ to 1.0	16	149.37	117.96
> 1.0+	28	232.48	415.56

* Market capitalization is in millions $$

Consistent with the findings of Knez and Ready (1997), Exhibit 7 reveals that the firms with the largest differences between the OLS and robust estimator (firms whose data most likely contain outliers) are on average the smallest firms in terms of market capitalization. In fact the average firm size and variation around that average decreases approximately monotonically as the absolute difference between the OLS and robust beta increases. The results in Exhibit 7 are qualitatively unchanged using the median and the median absolute deviation robust measure of scale indicating that this result is robust toward outliers in market capitalization.

Exhibit 8 provides a box plot Trellis display of the distributions of the logarithm of market capitalization, conditioned on the bin for absolute difference of OLS and robust beta, computing market capitalization in the same way as for Exhibit 7. The conditional box plots provide a clear picture of the distribution of market capitalization conditional on the difference between OLS and the robust estimator. The median size of the firm decreases as the difference between OLS and the robust estimator increases, i.e., as the influence of outliers on the OLS beta increases. Also, the dispersion of the market capitalization become more tightly clustered around the median as the influence of outliers on the OLS beta increases.

EXHIBIT 8
BOXPLOT VIEWS OF CONDITIONAL DISTRIBUTION OF LOG OF MARKET CAPITALIZATION, CONDITIONED ON ABSOLUTE DIFFERENCE BIN



6. CONCLUDING COMMENTS

Our comparison of the relative behavior of the classical OLS beta and the proposed robust beta for historical equity returns show emphatically that outliers can be a real cause for concern. Furthermore, the concern is primarily associated with small market capitalization firms.

Because a few outliers or even a single outlier can have such a strong distorting influence on the OLS beta, the OLS beta can provide an at worst misleading or at best incomplete indication of the risk versus expected return characteristics of a particular equity. If one were to rely on a single estimation method, the robust beta might well be preferred to the OLS beta. This is because: (a) the robust beta accurately reflects the risk versus expected return characteristics of the *bulk* of the returns for a given equity, without regard to the behavior of outliers, and (b) the bulk of the data is in our experience the only part of the data is well predicted with a probability distribution model. (Those who can predict outliers well should tell us how, so we can share in the profits!).

It is apparent that when influential outlier returns exist, neither the OLS beta nor the robust beta alone provide an adequate summary of an equities risk versus expected return characteristics. However, the two estimates might be used together with considerable effectiveness as follows. When the difference between the two estimates is less than a specified Δ, based on either a consumer's financial sensibility or a statistical significance test, report the classical result. Otherwise, report the value of both estimates, along with additional information such as the times of occurrence and sizes of outliers that are given zero weights by the robust regression estimate of beta. Desirable additional information would include event information associated with the times of occurrences of outliers, e.g., temporally nearby corporate announcements of potential relevance.

It is our recommendation that commercial suppliers of beta calculations provide a more complete picture to consumers by following some variant of the above suggestions.

BIBLIOGRAPHY

Becker, R. A. and W. S. Cleveland, “The Design and Control of Trellis Display,” *Journal of Computational and Statistical Graphics*, 5, (1996) 123-155.

Chan, L. K. C. and J. Lakonishok, “Robust Measurement of Beta Risk,” *Journal of Financial and Quantitative Analysis*, 27:2 (June 1992) 265-282.

Connolly, R. A., “An Examination of the Robustness of the Weekend Effect,” *Journal of Financial and Quantitative Analysis*, 24:2 (June 1989) 133-169.

Cook, R. D. and S. Weisberg, “Residuals and Influence in Regression,” Chapman and Hall, New York, (1982).

Cornell, B. and J. K. Dietrich, “Mean-Absolute-Deviation Versus Least-Squares Regression Estimation of Beta Coefficients,” *Journal of Financial and Quantitative Analysis*, 13 (March 1978) 123-131.

Fama, E. “The Behavior of Stock Prices,” *Journal of Business,* 38 (Jan. 1965) 34-105.

__________, and Kenneth R. French, “Cross-section of Expected Stock Returns,” *Journal of Finance*, 47 (1992) 427-466.

Hampel, F. R., E. M. Ronchetti, P. J. Rousseeuw, and W. A. Stahel, “Robust Statistics. The Approach Based on Influence Functions,” *Wiley Series in Probability and Mathematical Statistics: Probability and Mathematical Statistics*. John Wiley & Sons, Inc., New York, 1986.

Huber, P.J., “Robust Regression: Asymptotics, Conjectures and Monte Carlo,” *Annals of Statistics*, 1 (1973) 799-821.

__________, “Robust Statistics,” *Wiley Series in Probability and Mathematical Statistics*. JohnWiley & Sons, Inc., New York, 1981. (Reviewer: Rudolph J. Beran)

Judge, G. G., R. C. Hill, W. E. Griffiths, H. Lütkepohl, and T. C. Lee, “Introduction to the Theory and Practice of Econometrics,” Second edition. John Wiley & Sons, Inc., New York, 1988. (Reviewer: Kajal Lahiri)

Knez, P. J. and M. J. Ready (1997), “On the Robustness of Size and Book-to-Market in Cross-Sectional Regressions,” *Journal of Finance*; 52(4), September 1997, pages 1355-82.

Koenker, R., and G. Bassett, “Regression Quantiles,” *Econometrica*, 46 (1978) 33-50.

Kon, S. “Models of Stock Returns – A Comparison,” *Journal of Finance*, 39 (March 1984) 147-165.

Mandelbrot, B. “The Variation of Certain Speculative Prices,” *Journal of Business*, 36 (Oct. 1963) 394-419.

Martin, R. D., “Trellis Displays for Deep Understanding of Financial Data,” *Financial Engineering News*, 2:1 (1998).

Martin, R. D. and Simin, T., “Robust Estimation of Beta”, Technical Report No. 350, Dept. of Statistics, Univ. of Washington, March (1999).

Martin, R. D., V. J. Yohai, and R. H. Zamar “Min-Max Bias Robust Regression,” *The Annal of Statistics*, 17:4 (1989) 1608-1630.

Richardson, Matthew and Tom Smith, "A Direct Test of the Mixture of Distributions Hypothesis: Measuring the Daily Flow of Information," *Journal of Financial and Quantitative Analysis,* (1994) 101-116.

Roll, R., “R^2” *Journal of Finance*, 43 (July 1988) 541-566.

Sharpe, W. F., “Mean-Absolute-Deviation Characteristic Lines for Securities and Portfolios,” *Management Science*, 18:2 (Oct. 1971) B1-B13.

S-PLUS. User’s Guide and On-Line Documentation for S-PLUS 4.5, Data Analysis Products Division of MathSoft, Inc., Seattle, WA (1998).

Yohai, V. J., W. A. Stahel, and R. H. Zamar, “A Procedure for Robust Estimation and Inference in Linear Regression,” in *Directions in Robust Statistics and Diagnostics, Part II*, Stahel, W. A. and S. W. Weisberg, Eds., Springer-Verlag, New York (1991).

Yohai, V. J., and R. H. Zamar, “Optimal Locally robust M-estimates of Regression,” *Jour. Statist. Inference and Planning* (1997).

APPENDIX

The Analytic Expressions for the Weight Function *w*

$$w(r;c)=\begin{cases} 0 & |r/c|>3 \\ \left[-1.944\cdot+1.728\cdot\left(\frac{r}{c}\right)^2-0.312\cdot\left(\frac{r}{c}\right)^4+0.016\cdot\left(\frac{r}{c}\right)^6\right], & 2<|r/c|\le 3 \\ r & |r/c|\le 2 \end{cases}$$

The function ψ can be obtained from w using (7), and then ρ can be obtained by integration.

Endnotes

[1] Our personal experience is that Gaussian mixture models often provide quite good approximations for outlier-generating conditional distributions of individual firm equity returns given market returns.

[2] Outliers generated by symmetric non-Gaussian distributions can also result in quite low *efficiency*. Efficiency is the ratio of the minimum attainable variance over all estimators to the actual variance of an estimate. OLS has an efficiency of one when linear model error term is Gaussian, but can have arbitrarily low efficiency for non-Gaussian error distributions. It is to be noted that even when the errors are non-Gaussian, the OLS estimates are still unbiased under fairly general conditions. However this is not the case when outliers in returns are caused by special firm-specific asymmetric events that do not conform to a symmetric error distribution in the market model (1). Such outliers can cause *bias*, even in large samples.

[3] See for example the books by Huber (1981) and Hampel et. al. (1986), and the references therein. A major goal of modern robust regression methods is to achieve small bias for arbitrary outlier contamination, while at the same time achieving a high efficiency, e.g., 90% when the data is Gaussian. See for example Hampel et. al. (1986), Martin, Yohai and Zamar (1989), and Yohai and Zamar (1997).

[4] In economics and finance, the least absolute deviations (LAD) estimate is perhaps the oldest and most widely known (somewhat) robust alternative to least squares. Sharpe (1971) studied LAD beat estimates for thirty common stocks used to compute the Dow Jones Industrial Average, and for thirty mutual funds, both in the mid-to-late 1960's. Somewhat later Cornell and Dietrich (1978) also studied the LAD estimate using one hundred companies randomly drawn from the S&P 500 from 1962 to 1975. Both Sharpe and Cornell and Dietrich concluded that the LAD alternative did little to improve the OLS estimate of beta. These negative results are no doubt due to the lack of sufficiently large and influential outliers in the returns of mutual funds and in the relatively large sized firms considered by these authors. This is quite the opposite of the current-day situation for small-sized firms where the returns often contain very large highly influential outliers. Subsequent positive results bhy Koenker and Basset (1978), Connolly (1989), and Chan and Lakonishok (1992) provided evidence that robust alternatives to OLS could be preferable for dealing with outlier-generating heavy-tailed returns distributions.

[5] It may be noted that in the regression context the influence of an outlier can be particularly potent when it is a *leverage* point, namely when it is associated with an outlying value of the predictor variable(s). In the context of the market model (1), this occurs when a particular market return is an outlier relative to the bulk of the market returns. Further discussion of outliers with leverage may be found for example in Judge et. al. (1988) and Cook and Weisberg (1982).

[6] There is another, perhaps more subtle effect of outliers on standard error estimates, namely leverage points can *reduce* the estimated standard error estimate. This occurs because leverage points have a large positive influence on the values of the sums-of-squares and cross products of the predictor variables, and an increase in these values results in an increase in the stated precision of parameter estimation. This second effect is evident in the increase in the standard errors of the OLS estimates for the four firms in Exhibit 2 when the October 1987 leverage point deleted (compare Panels A and B in Exhibit 2). It can be shown that a leverage point influences the standard error estimate for our robust beta estimate only if the residual for the leverage point is small.

[7] A test statistic for testing for bias in the OLS estimate is discussed in Martin and Simin (1999), who provide results of using the test statistic on the estimate reported in the present paper.

[8] Regression M-estimates were first introduced by Huber (1973), as a generalization of his M-estimates of location (Huber, 1964). See also, Huber (1981) and Hampel et. al. (1986).

[9] The ordinary least squares (OLS) loss function is $\rho_{OLS}(r) = r^2$ and the least absolute deviation (LAD) loss function is $\rho_{LAD}(r) = |r|$. In these cases the scale estimate $\hat{s}$ is not needed.

[10] Huber's favorite loss function ρ was the *unbounded* convex function that is quadratic (like OLS) in a central interval (-*c*, *c*) and linear (like LAD) outside that interval, and consequently is a compromise between the OLS and LAD estimates. However, it can be shown that any unbounded loss function can result in arbitrarily large parameter estimate bias due to mixture model outliers, and that bounded loss functions are needed in order to obtain good bias robustness properties (Martin, Yohai and Zamar, 1989).

[11] The box plots are computed for the OLS and robust beta estimates conditional on the absolute difference being in a given bin. The solid dot represents the median of the distribution. The box indicates the inter-quartile range (IQR: the middle half of the data), the whiskers cover the range 1.5* IQR and points beyond the whiskers may be considered outliers.

[12] For general information on Trellis displays, see Becker and Cleveland (1998), and for applications to analyzing equity returns, see Martin (1998). We used the Trellis display software contained in S-PLUS (1998) to make the display of Exhibit 5 and also that of Exhibit 7.

Examples of outlier in a sentence, how to use it. 99 examples: Outliers that arise simply from the skewness of the distribution can beâ€¦Â The 10 high reactive outliers, compared with the 55 remaining high reactives, were either very quiet or very talkative at 41/2 years. From Cambridge English Corpus. No other variables had outliers that were found to skew the results.