BICKEL, P. J. and LEHMANN, E. L. (1975). Descriptive statistics for non-parametric models. II. *Ann. Statist.* **3** 1045–1069.

BROWNLEE, K. A. (1965). *Statistical Theory and Methodology in Science and Engineering*, 2nd ed. Wiley, New York.

CARROLL, R. J. (1979). On estimating variances of robust estimators when the errors are asymmetric. *J. Amer. Statist. Assoc.* **74** 674–679.

DE WET, T. and VENTER, J. H. (1974). An asymptotic representation of trimmed means with applications. *S. African Statist. J.* **8** 127–134.

HILL, M. A. and DIXON, W. J. (1982). Robustness in real life: a study of clinical laboratory data. *Biometrics* **38** 377–396.

HUBER, P. J. (1972). Robust statistics: a review. *Ann. Math. Statist.* **43** 1041–1067.

HUBER, P. J. (1981). *Robust Statistics*. Wiley, New York.

JAECKEL, L. A. (1971). Some flexible estimates of location. *Ann. Math. Statist.* **42** 1540–1552.

KOENKER, R. W. and BASSETT, G. W. (1978). Regression quantiles. *Econometrica* **46** 33–50.

KOUL, H. L. (1969). Asymptotic behavior of Wilcoxon type confidence regions in multiple linear regression. *Ann. Math. Statist.* **40** 1950–1979.

ROCKE, D. M., DOWNS, G. W. and ROCKE, A. J. (1982). Are robust estimators really necessary? *Technometrics* **24** 95–101.

RUPPERT, D. and CARROLL, R. J. (1980). Trimmed least squares estimation in the linear model. *J. Amer. Statist. Assoc.* **75** 828–838.

SHORACK, G. R. (1974). Random means. *Ann. Statist.* **2** 661–675.

SKOROKHOD, A. V. (1956). Limit theorems for stochastic processes. *Theory Probab. Appl.* **1** 261–290.

SPJØTVOLL, E. and AASTREIT, A. H. (1980). Comparison of robust estimators on data from field experiments. *Scand. J. Statist.* **7** 1–13.

STIGLER, S. M. (1973). The asymptotic distribution of the trimmed mean. *Ann. Statist.* **1** 472–477.

STIGLER, S. M. (1977). Do robust estimators work with real data? *Ann. Statist.* **5** 1055–1077.

TUKEY, J. W. and MCLAUGHLIN, G. (1963). Less vulnerable confidence and significance procedures for location based on a single sample. *Sankyā Ser. A* **35** 331–352.

DEPARTMENT OF STATISTICS
UNIVERSITY OF CHICAGO
5734 UNIVERSITY AVENUE
CHICAGO, ILLINOIS 60637

DISCUSSION

P. J. DE JONGH AND T. DE WET

Institute for Maritime Technology

This paper proposes a new way of defining trimmed means in the linear model, which differs from earlier proposals by Bickel (1973), Koenker and Bassett (1978) and Ruppert and Carroll (1980). We find the idea of the proposal very interesting. It has the "right" equivariance and asymptotic properties and is thus an attractive (large sample) extension of the trimmed mean in the location case. These properties also hold for the Koenker–Bassett (1978) estimator, but the Welsh estimator has the potential advantage of computational simplicity (if least squares is used as a preliminary estimator). Our remarks will concern the small sample behaviour of the proposed estimator. We wish to

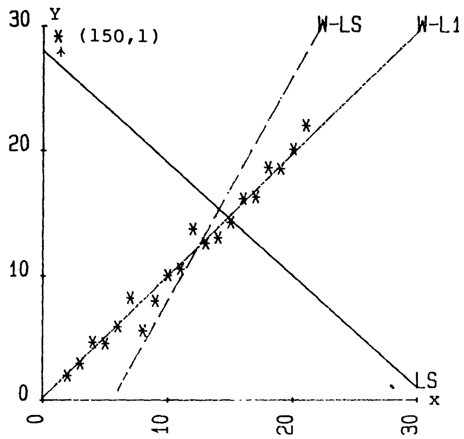

FIG. 1. *One y outlier.*

illustrate this with a few simple examples. Rather extreme examples were chosen for illustrative purposes. For less extreme cases, similar results hold albeit not as pronounced.

Our first point concerns the role of the initial estimator. Consider the model $Y_i = i + e_i$, with $\{e_i\}$ iid standard normal. From this model $n = 21$ points were generated and point 1 is "moved out" in the y direction to become an outlier. With this y value equal to 150, the 10% Welsh trimmed mean with least squares (LS) as initial estimator was calculated and the fitted line is denoted in Figure 1 by W-LS.

In this figure we also indicate the LS line as well as W-L1, the 10% Welsh estimator with L_1 as preliminary. Note that the LS line is pulled toward the outlier, as one would expect, creating large positive as well as negative residuals on the edges of the design space. Thus, in this case of an outlier on the edge of the design space, points on the two edges are overadjusted for leading to the W-LS line overestimating the slope. Clearly this is caused by the fact that LS is used as the initial estimator. Using, e.g., L_1 as initial estimator, we do not have this problem and W-L1 goes through the bulk of the points. Thus, although W-LS has a computational advantage over W-L1, the latter might still be preferable from a resistance point of view. Note that the 15% Koenker–Bassett trimmed mean is of the same order of computational difficulty as W-L1 and is also not sensitive to this type of outlier. Note also that outliers in the centre of the design space are handled well by W-LS, e.g., if point 11 is moved out in the y direction.

Our second point concerns the effect that outliers in the design space have on the Welsh estimator. Although it was not specifically designed to cater to this situation, we feel it is worth discussing. Consider again the above model, but now move point 1 out in the x direction to have an x value of 150.

In Figure 2 we again give W-LS, W-L1 and LS. In this case, both W-LS and W-L1 do not improve on the disastrous behaviour of LS. This is a typical

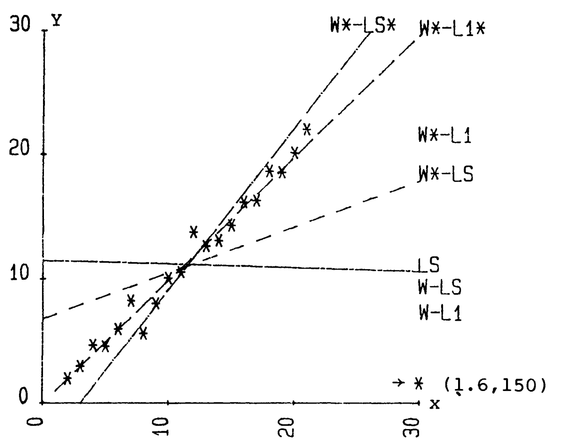

FIG. 2. *One x outlier.*

situation where a bounded influence-type estimator is needed (see e.g., Krasker and Welsch (1982)). The following seems a natural way of achieving this for the Welsh estimator. Let $\{w_i\}$ be a set of weights, typically dependent on $\{x_i\}$ and in the definition of τ_n replace x_i by $w_i x_i$. Taking $w_i = u(x_i)$ for some function u, it follows, along the lines of Welsh's proofs, that under appropriate regularity conditions, the resulting estimator has influence function of the form

$$Q_u^{-1} x u(x) \phi(y - x'\theta_0)$$

for a certain matrix Q_u (we considered the case of a symmetric F and symmetric trimming for simplicity). Choosing an appropriate u function then gives a bounded influence function. However, in small samples, this estimator is again adversely affected by the choice of initial estimator if the latter does not have bounded influence. We thus propose the following. Modify the initial estimator to have bounded influence and then use this in the above bounded influence Welsh estimator. Denote by W*-LS and W*-LS*, respectively, the above bounded influence estimators with LS and weighted LS as preliminary estimators. Similarly we use the notation W*-L1 and W-*L1*. Note that in this case L1* is a bounded influence L_1 estimator as defined in de Jongh and de Wet (1985). For $\{w_i\}$ we use the Mallows weights (see, e.g., Denby and Larsen (1977)) with 15% trimming. The results are also given in Figure 2. We see that in this case W*-LS* does reasonably well, while W*-L1* does very well. In Figure 3 we considered the case of a y and x outlier. All the previous estimators are again shown and clearly only W*-L1* goes through the bulk of the data.

If the above estimators are applied to the salinity data (which has two clear x outliers, viz. points 5 and 16) it is seen that W*-L1* gives the best fit in terms of IQR. This makes one fairly confident about the robustness of W*-L1* to x and y outliers. We also note that the bounded influence Koenker–Bassett trimmed means defined in de Jongh and de Wet (1985) have similar behaviour to W*-L1*.

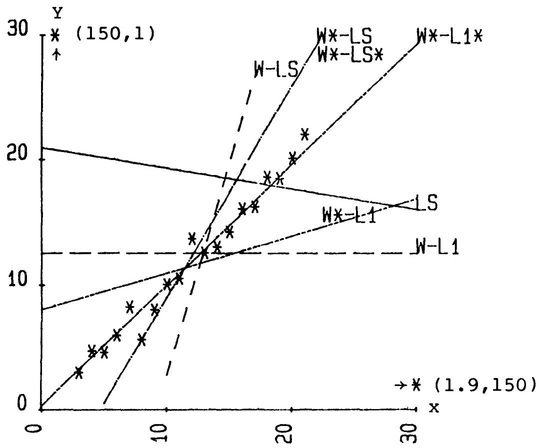

FIG. 3. *One y and one x outlier.*

Although our results are based on a limited number of data sets, we conclude that the Welsh trimmed mean should be used with caution in practice.

REFERENCES

BICKEL, P. J. (1973). On some analogues to linear combinations of order statistics in the linear model. *Ann. Statist.* **1** 597–616.

DE JONGH, P. J. and DE WET, T. (1985). Trimmed means and bounded influence estimators for the parameters of the AR(1) process. *Comm. Statist. A—Theory Methods* **14** 1361–1375.

DENBY, L. and LARSEN, W. A. (1977). Robust regression estimators compared via Monte Carlo. *Comm. Statist. Theory Methods* **6** 335–362.

KOENKER, R. W. and BASSETT, G. W. (1978). Regression quantiles. *Econometrica* **46** 33–50.

KRASKER, W. S. and WELSCH, R. E. (1982). Efficient bounded-influence regression estimation. *J. Amer. Statist. Assoc.* **77** 595–604.

RUPPERT, D. and CARROLL, R. J. (1980). Trimmed least squares estimation in the linear model. *J. Amer. Statist. Assoc.* **75** 828–838.

INSTITUTE FOR MARITIME TECHNOLOGY
P.O. BOX 181
SIMON'S TOWN 7995
REPUBLIC OF SOUTH AFRICA

ROGER KOENKER

University of Illinois, Urbana

1. Introduction. Alan Welsh has resolved an intriguing puzzle posed by Ruppert and Carroll (1980) in their influential study of analogues of the trimmed mean for the linear regression model. They showed that an estimator with "appropriate" asymptotic behavior could be constructed based on "regression quantiles," and they also showed that naive trimming based on residuals from a preliminary fit of the model had a considerably different, and far less satisfactory, asymptotic theory. Welsh has now shown that a less naive, but still