At a basic level, there is a question as to whether, in Draper's notation, we should regard $1/\theta$ or θ as the parameter to be estimated. I have a personal preference for $1/\theta$ because this seems more natural to me and because I feel that the bias enters in a simpler way than when we take the reciprocal of an estimator of θ.

For the L_1-estimator (which is related to the R-estimator with sign scores; see (3.10) in Draper), we can construct a kernel estimator of $1/\theta$ directly (Welsh, 1987c). What is interesting about this estimator is that the shape of the kernel or window function does seem to matter as a poor choice can lead to an estimator with excessive bias. This is in conflict with the usual advice (reported by Draper) that in estimating a density, the choice of kernel is unimportant.

In evaluating competing estimates of the variance of an R-estimator, we should evaluate their properties as studentizing factors rather than as estimates of the variance per se. Although this is quite often done in simulation studies, it is not often done in theoretical investigations. However, recently Hall and Sheather (1988) derived an Edgeworth expansion for the sample median studentized by a particular variance estimator and showed that the optimal choice of smoothing parameter is different from that obtained from mean squared error considerations. In fact, their result indicated that it is important to decrease the bias more than one would if the variance was a parameter of interest. In other words, the bias/variance tradeoff is different when the density is a nuisance parameter than when it is a parameter of interest. These results are in agreement with the practical experience reported by Draper that the bias is more important than the variance in estimating $1/\theta$ (or θ).

4. *L*-, *M*- OR *R*-ESTIMATORS?

In advocating the use of R-estimators over M-estimators, Draper notes only that they often have simple, closed-form expressions. He does not mention that perhaps a more serious objection to M-estimation is that scale equivariance is usually achieved through the use of a concomitant scale estimator which may have subtle effects on the properties of the M-estimator and on the resulting inference. Now L-estimators (Welsh, 1987b; Koenker and Portnoy, 1987) have been developed further since Draper's work and they share the advantages of R-estimators. However, they have one further advantage: if the weight function is chosen to be smooth, the asymptotic variance of the resulting L-estimator is straightforward to estimate. That is, the complete analysis (including inference) is easier for L-estimators than for R-estimators. Consequently, I welcome Draper's paper for the indirect support it provides for the use of L-estimators in the linear model problem.

ADDITIONAL REFERENCES

Carroll, R. J. and Welsh, A. H. (1987). A note on asymmetry and robustness in linear regression. *Amer. Statist.* To appear.

Hall, P. and Sheather, S. J. (1988). On the distribution of a studentised quantile. *J. Roy. Statist. Soc. Ser. B.* To appear.

Hampel, F. R., Ronchetti, E. M., Rousseeuw, P. J. and Stahel, W. A. (1986). *Robust Statistics: The Approach Based on Influence Functions.* Wiley, New York.

Koenker, R. and Portnoy, S. (1987). L-estimation for linear models. *J. Amer. Statist. Assoc.* **82** 851–857.

Welsh, A. H. (1987b). One-step L-estimators for the linear model. *Ann. Statist.* **15** 626–641.

Welsh, A. H. (1987c). Kernel estimates of the sparsity function. In *Statistical Data Analysis Based on the L_1-Norm and Related Methods* (Y. Dodge, ed.) 369–378. North-Holland, Amsterdam.

Comment

Roger Koenker and Stephen Portnoy

David Draper's survey of rank-based robust methods for estimation and inference in linear models vividly illustrates the vitality of the R approach. The emphasis on inference is, in our view, particularly welcome, because despite the rapid growth of the foundations of robust *estimation* for linear models, the framework for robust *inference* has languished in a state of benign neglect. Certainly in applied fields like econometrics, unless we are able to suggest simple, yet reliable, robust methods of computing "those little numbers in parentheses," robust methods in general will continue to be a curiosity of the "theorists" with little impact on empirical research.

On Draper's three *desiderata* for a successful robust method: (i) intuitive appeal, (ii) unified theory and (iii) computability, we would like to offer some highly

Roger Koenker is Professor of Economics, Department of Economics and Stephen Portnoy is Professor of Statistics, Department of Statistics, University of Illinois, Champaign, Illinois 61820.