context, this paper does not provide the reader with the necessary ideas. Although many papers have been published in this area since 1982, only two papers published since then are cited. Discussion of many of the fundamental issues, including a comprehensive review of the literature to 1982, can be found in Cook and Weisberg (1982, especially Section 5.2); see also Weisberg (1983) and Cook (1986).

ADDITIONAL REFERENCES

COOK, R. D. (1986). Assessment of local influence (with discussion). To appear in *J. Roy. Statist. Soc. Ser. B* **48**.

WEISBERG, S. (1983). Some principles for regression diagnostics and influence analysis, discussion of a paper by R. R. Hocking. *Technometrics* **25** 240–244.

WEISBERG, S. (1985). *Applied Linear Regression*, 2nd ed. Wiley, New York.

Rejoinder

Samprit Chatterjee and Ali S. Hadi

Many points have been raised, but alas, space does not permit us to respond to each one of them individually. For expedience, the comments which we feel have arisen due to a misreading of what we wrote will not be discussed, letting the readers make up their own minds. Our paper will have served its purpose if it stimulates discussion and leads to further development in methodology. We are grateful to Professor DeGroot for getting together such a distinguished group to act as discussants for our paper.

Several of the authors (Brant, Hoaglin and Kempthorne, and Welsch), have pointed out very correctly that little was said in our paper about detecting groups or clusters of influential points. Not much is known; and we came to know about the work of Brant and Kempthorne only recently. We are not convinced, however, as to how real the problem is. Most of the influential points may be detected by a one point at a time deletion scheme. We see our skepticism on this point is also shared by Welsch.

Weisberg has noted that we have not provided an overriding general principle for deriving various influence measures. Space considerations prevented such an effort. Basically we tend to favor the influence function approach introduced by Hampel. In our forthcoming book, *Sensitivity Analysis in Linear Regression*, we outline such an approach. We show that almost all proposed influence measures can be derived from various approximations of the influence function. The likelihood approach, as pointed out by Cook and Weisberg, is another unifying principle. We are not convinced, however, about its robustness. We prefer measures which are based on metric distances rather than those based on probability densities, and therefore we have stayed away from influence measures based on information theory.

Several authors have raised questions about the callibration points which we have provided in Table 2. There is nothing sacred about them. They are meant to be yardsticks, equivalent to "± 2 standard error rules." Our attitude to them is identical with those articulated by Velleman and Hoaglin and Kempthorne, although it might not have been stated as explicitly. Points which stand out from the group on their diagnostic measures should certainly be flagged and examined. It is the standing apart which should trigger off the alarm rather than the exceeding of a critical value. Stem and leaf plots are very effective graphical devices for this purpose. We would like to endorse the diagnostic strategy advocated by Hoaglin and Kempthorne. In fact it is this approach which has led us to flag points 1 and 17 on the basis of CVR_i rather than all the points which mechanically lie outside the critical interval. Points in Table 5 are starred only when they stand out (outliers on the diagnostic measure) rather than merely exceed their calibration values. We thank the discussants for highlighting this point.

Several of the discussants brought up the important question of observations influencing variable selection in model determination. Most influence measures do not distinguish whether an observation is influential on all dimensions or only on one or few dimensions. An observation, for example, might appear to be the most influential one according to a given measure, but when a particular variable is omitted the influence disappears. Retaining a variable may hinge on one or a few observations. In our present paper, we did not discuss this complex question, but have a paper languishing somewhere in the refereeing process, which addresses this question. The role of observations in variable selection (irrespective of the criteria used) is an area which needs clarification.

Atkinson in his related comments makes a point not made by the other discussants. If we have read his comments correctly, it appears that he opts for a robust estimation procedure. This is certainly a valid approach. A model fitting approach in which no point has an excessive influence certainly gets rid of the problem of influential points. We feel, however, that such a mechanical approach misses the creative aspect of data analysis. It sweeps a lot of problems under the rug. The diagnostic approach will reveal features which would be missed in a mechanical robust fitting. An example in point is the Moore data, which has now been described in detail by Weisberg. All the diagnostic measures point up observation 17; this is now acknowledged to have a transcription error. We regard this identification as a confirmation, if that was needed, of the value of the diagnostic measures.

The diagnostic measures that we have presented are useful and should play an important part in data analysis. But as Velleman points out their use will become widespread only if commonly available statistical software implements them. Let us hope that this is forthcoming, and we hope people like Velleman and Welsch will take the lead in it. Before expert systems and smart software take over we must agree on what the most effective approach is, or else we will be implementing mechanically rigid procedures like the step-wise methods for variable selection.

The last question which we take up is the question of notation and terminology. On this point we have apparently stepped on several toes. Hoaglin and Kempthorne's plea, "A consensus on notation for the basic quantities in regression diagnostics would be most welcome," should be heeded. We thought that we were attempting a step in that direction. Let us explain: Consider the two matrices

$$P = X(X^TX)^{-1}X^T \qquad \text{and} \qquad R = (I - P)$$

which occur extensively in linear regression analysis. We called them the prediction (projection) matrix and the residual matrix because applying them on Y produces the predicted values and the residuals, respectively. Prediction and projection are more widely understood operations (although less colorful) than capping or "hatting." The Hat (Hoaglin?) matrix leaves almost all first-time listeners mystified! Belsley, Kuh, and Welsch's book, *Regression Diagnostics*, was a very valuable contribution to the statistical literature, but it unleashed on an unsuspecting statistical community a computer-speak (à la Orwell) the likes of which we have never seen. We aesthetically rebel against DFFIT, DFBETA, etc., and have attempted to replace them by the last name of the authors according to a venerable statistical tradition. We hope that this approach proves attractive to the statistical community. Only time will tell!

We conclude by thanking all the discussants for their valuable comments. They were stimulating, interesting, and we hope will lead to more work in this area. We take heart from a comment by Wittgenstein in his *Tractatus*, "We can make nothing clear, but only some things clearer."