5. *Study Influence of Individual Observations on Fit.* We customarily plot DFITS_i (WK_i) against i.

6. *Study Influence of Individual Observations on Estimates of Coefficients.* For each j we plot $\text{DBETAS}_{ij}(D^*_{ij})$ against i, and we look at these plots in parallel.

7. *Study Influence of Individual Observations on the Estimated Covariance Matrix of* $\hat{\beta}$. Here we plot COVRATIO_i (CVR_i) against i. In Steps 5 and 7 we also examine the residual versus leverage plots with iso-influence contours.

8. *Probe for Subsets of Observations That Are Jointly Influential.* Although more research is needed in this area, we feel it forms an important part of the diagnostic strategy. The k-clustering approach of Gray and Ling (1984) and the derivative influence techniques of Kempthorne (1986) seem promising. Another, more ad hoc, approach is to drop the observations (say, three or four) that have the most individual influence and then see how much the results change.

For a diagnostic analysis, this strategy constitutes a bare minimum. Often, other areas of diagnosis are critical to the analysis: need for transformation, influence on model choice, or detecting departures from the standard Gauss-Markoff assumptions such as heteroscedastic or correlated errors. Research in these areas among others has been especially active in recent years, including applications of a Bayesian perspective. See, e.g., Atkinson (1982), Cook and Weisberg (1983), Dawson (1985), Johnson and Geisser (1983), and Pettit and Smith (1985).

ACKNOWLEDGMENTS

This work was supported in part by Contract DAAG29-85-K-0262 between the United States Army Research Office and Harvard University and by National Science Foundation Grant SES-8401422.

ADDITIONAL REFERENCES

ATKINSON, A. C. (1978). Posterior probabilities for choosing a regression model. *Biometrika* **65** 39–48.

ATKINSON, A. C. (1981a). Likelihood ratios, posterior odds, and information criteria. *J. Econometrics* **16** 15–20.

CLEVELAND, W. S. (1985). *The Elements of Graphing Data.* Wadsworth, Monterey, Calif.

COOK, R. D. and WEISBERG, S. (1983). Diagnostics for heteroscedasticity in regression. *Biometrika* **70** 1–10.

DAWSON, R. (1985). Diagnosing data and prior influence in a Bayesian analysis. Unpublished Ph.D. thesis, Dept. Statistics, Harvard Univ.

GRAY, J. B. (1983). The L − R plot: a graphical tool for assessing influence. In *1983 Proceedings of the Statistical Computing Section* 159–164. Amer. Statist. Assoc., Washington, D. C.

GRAY, J. B. (1985). Graphics for regression diagnostics. In *1985 Proceedings of the Statistical Computing Section* 102–107. Amer. Statist. Assoc., Washington, D. C.

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

JOHNSON, W. and GEISSER, S. (1983). A predictive view of the detection and characterization of influential observations in regression analysis. *J. Amer. Statist. Assoc.* **78** 137–144.

JOINER, B. L. (1981). Lurking variables: some examples. *Amer. Statist.* **35** 227–233.

KEMPTHORNE, P. J. (1985). Decision-theoretic measures of influence in regression. In *1985 Proceedings of the Business and Economic Statistics Section* 429–434. Amer. Statist. Assoc., Washington, D. C. To appear in *J. Roy. Statist. Soc. Ser. B.*

KEMPTHORNE, P. J. (1986). Identifying derivative-influential groups of observations in regression. Memorandum NS-540, Dept. Statistics, Harvard Univ.

KRASKER, W. S. and WELSCH, R. E. (1983). The use of bounded-influence regression in data analysis: Theory, computation, and graphics. In *Computer Science and Statistics: Proceedings of the 14th Symposium on the Interface* (K. W. Heiner, et al., eds.) 45–51. Springer, New York.

PETTIT, L. I. and SMITH, A. F. M. (1985). Outliers and influential observations in linear models. In *Bayesian Statistics* **2** (J. M. Bernardo, M. H. DeGroot, D. V. Lindley, and A. F. M. Smith, eds.) 473–494. North Holland, Amsterdam.

SCHWARZ, G. (1978). Estimating the dimension of a model. *Ann. Statist.* **6** 461–464.

WELSCH, R. E. (1983). Discussion of Developments in linear regression methodology: 1959–1982, by R. R. Hocking. *Technometrics* **25** 245–246.

Comment

Paul F. Velleman

I congratulate Chatterjee and Hadi on an excellent survey of an area that has developed rapidly in the past decade. One of the disappointments of this area is that these very valuable techniques have been slow to infiltrate the literature of disciplines using regression techniques. We need to turn some of our attention to promoting the use of diagnostic statistics in ordinary practical analyses.

One problem with regression diagnostics has been that terminology has not yet standardized. Unfortunately, Chatterjee and Hadi exacerbate rather than alleviate this problem. I do not believe that we need yet another name and notation for the Hat matrix, nor that we benefit from new and somewhat cryptic

Paul F. Velleman is Associate Professor of Economic and Social Statistics, Cornell University, 358 Ives Hall, Ithaca, New York 14853.