Regression Diagnostics:
Robust *versus* Least Squares Residuals

by

R. Dennis Cook and Sanford Weisberg

University of Minnesota

School of Statistics

Department of Applied Statistics

St. Paul, Minnesota 55108

Technical Report Number 538

January 1990

Regression Diagnostics: Robust *versus* Least Squares Residuals

R. D. Cook and S. Weisberg *

January 1990

Abstract

Several well known examples are reworked to compare diagnostics produced using ordinary least squares and least median of squares estimates.

1 Introduction

Robust estimation and regression diagnostics are often thought of as complementary concepts. Underlying both, one finds a concern for the possibility of incorrect assumptions in a model, incorrect data, overly influential data, or some combination of these. Robust estimates are designed to give reasonable estimates for a range of possible assumptions; most typically robust estimators are designed to perform well when the error distributions are not the ones usually assumed. Regression diagnostics take a somewhat different view. One fits using standard techniques and

*Department of Applied Statistics, University of Minnesota, St. Paul, Minnesota 55108. We are indebted to the stimulating discussions of many of the participants in the workshop on robustness and diagnostics at the Institute for Mathematics and Its Applications of the University of Minnesota for motivating our working these examples. We thank Douglas Hawkins for helping with the computations on two of the examples, and Raymond Carroll for many stimulating discussions. This work was supported by the Natinal Science Foundation, grant number DMS-8801544-02.

then, conditioning on the results of the standard fit, one then attempts to discover any incorrect assumptions. The area of greatest overlap between robust estimation and regression diagnostics seems to be in problems where outliers might be present. Robust methods *accomodate* outliers; regression diagnostics *identify* them.

During the summer of 1989, the Institute for Mathematics and its Applications at the Unversity of Minnesota hosted a workshop on robust estimation and diagnostics. One of the continuing themes during the sessions was on the complementary uses of robust estimation and diagnostic methodology. Several diametrically opposed points of view were expressed on this issue, which can be partly summarized as follows:

- Robust estimation methods render diagnostic methods generally unnecessary since, if the breakdown is high enough, the robust method can accommodate a proportion of "bad" data. Besides, results from a robust fit, such as the residuals and the estimated weights, provide any needed diagnostics (see, for example, Krasker and Welsch, 1982, page 602). Similarly, Hampel, *et. al* (1986, p. 12) wrote "...the residuals from a robust fit automatically show outliers and the proper random variability of the 'good' data, much clearer than for example residuals from least squares ..." In Rousseeuw and Leroy (1987, p. 92-3) we find "...residuals associated with a robust fit yield powerful diagnostic tools. For example, they can be displayed in residual plots. These graphs make it easy to detect outlying values and call attention to model failures."

- Doing diagnostics is always important, but why not start with a robust estimator, and then apply usual diagnostics methods to the results of the robust fit? In particular, for the purpose of assessing model adequacy and after ignoring any outlying points that have been heavily downweighted by the robust fitting procedure, plots of residuals from the robust fit can be interpreted as plots of residuals from a least squares fit and provide all necessary diagnostic information.

- In the context a linear regression model $Y = X\beta + \epsilon$, robustness makes getting the stochastic part of the model (ϵ) right unnecessary but provides little help in the diagnostic criticism of the functional part of the model ($E(Y) = X\beta$). Diagnostics, on the other hand, help to get the functional part right but provide little help for the stochastic part.

- Compared to standard likelihood inference, the robust estimation paradigm is too poorly understood to provide a basis for doing model/data diagnostics. In particular, least squares is a useful basis for diagnostics in linear models precisely because it is well understood and is very sensitive to departures from the model.

Our goal in this paper is to study several well known examples from the literature to address the relevance of the above points of view. We chose several linear regression examples that we have studied extensively. We fit these data by using the implementation of least median of squares (LMS) regression in the *Progress* program of Rousseeuw and Leroy (1987), and compare the results to those based on other diagnostic methods. LMS and related methodology was chosen because its breakdown point is essentially 50% and because it is one of the more prominent robust methods currently under development. Results for other robust methods are mentioned in Section 3. We will present no new theory, but merely attempt to highlight differnces between the two approaches, and perhaps point out new directions of research for those interested in promoting the robust estimation paradigm.

In all the examples in this paper, we number cases starting with zero, rather than with one, as is done in *lisp*. Also, we use the rule that a case will be called an LMS outlier if *Progress* gives that case weight zero when fitting reweighted least squares. *Progress* was used to do least median of squares calculations. A *FORTRAN* program written by Douglas Hawkins was used to calculate the minimum volume ellipsoids in Examples 2.1 and 2.4. Other robust estimators were computed using *New S*. All remaining calculations and all the graphs were done using *XLISP-STAT*.

2 Examples

2.1 Rat data

Reference: Weisberg (1985, p. 122). Nineteen rats were weighed, given a dose of a drug roughly proportional to body weight and, after a time, sacrificed. The response is the proportion of the drug recovered from the liver and the explanatory variables are liver weight, body weight and dose. The experimenter thought that the response should be independent of the three predictors, but this does not appear to be so. In the usual analysis, one finds that case 2 is influential, which is easily discovered by using Cook's distance, and any significance in the regressors is due to this one case. An elegant diagnostic for finding this is given by Cook and Weisberg (1989, p. 280). Consultation with the investigator revealed that a serious error was made in the experimental protocol for case 2. We regard any analysis that does not identify case 2 as deficient.

Figure 1 gives the OLS plot of residuals versus fitted values. This plot shows case 2 separated from the others, but with a relatively small residual. That this is a symptom of an influential case can be seen by using the local influence methods discussed in Section 1.4.2 of Cook and Weisberg (1990). Figure 2 is the plot of LMS residuals versus LMS fitted values. LMS finds two cases, 0 and 18, to be outliers. Case 2 is included in the elemental subset used to compute the LMS estimates, and thus case 2 has residual zero. The residual plot has an unusual and unexplained shape, more or less an x-pattern. Evidently, no theory is available to tell if such a pattern is meaningful.

The theory of least squares in combination with Figure 1 indicates the need for special attention to case 2; indeed, if case 2 is deleted the dimension of the

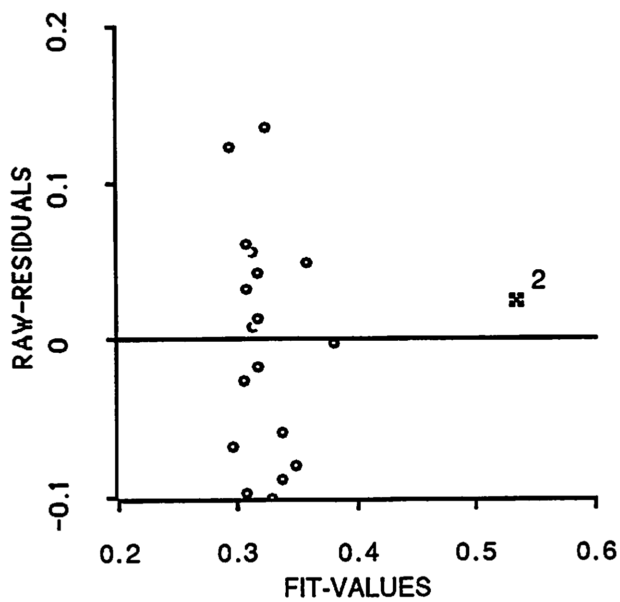


Figure 1: Rat data: OLS residuals versus fitted values.

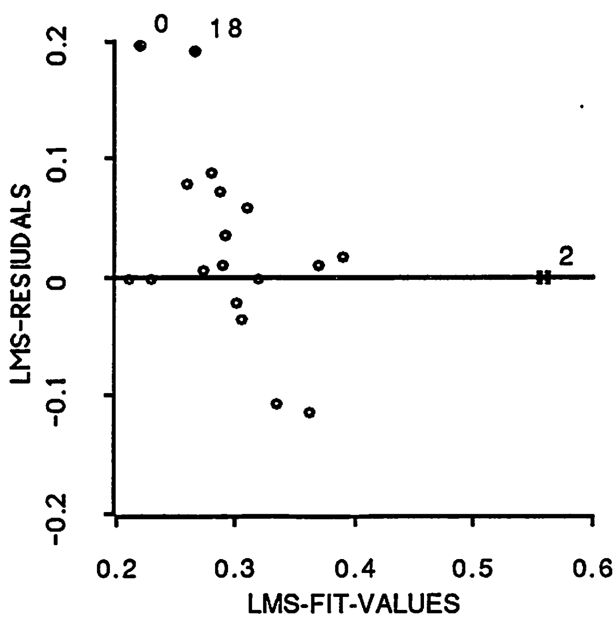


Figure 2: Rat data: LMS residuals versus fitted values. Cases 0 and 18 are LMS outliers.

estimation space is essentially reduced. If one can apply least squares theory to LMS regression, which is not a settled question, then case 2 is also found by LMS; however, it is not found by looking at the robust weights or the robust residuals. In addition, LMS labels as outliers two points that are more than likely good data.

Case 2 is influential primarily because it is outlying in the factor space: the corresponding element of the hat matrix $H = X(X^TX)^{-1}X^T = (h_{ij})$ is $h_{22} = 0.85$. Rousseeuw and van Zomeren (1990) recommend that leverage be assessed by using the robust Mahalanobis distances RD_i based on the inner product that comes from the minimum volume ellipsoid (MVE). This is defined to be the ellipsoid with smallest volume that contains a specified fraction of the data, or, equivalently, by finding the smallest volume ellipsoid covering c of the data points. Rousseeuw and Leroy (1987, p. 264) suggest setting $c = [(n+p+1)/2]$, which is $[(19+3+1)/2] = 11$ for this example. Rousseeuw and van Zomeren (1990) provide a method for calibrating the robust Mahalanobis distances to declare cases to be outliers in the factor space and thus leverage points. Table 1 summarizes the results of applying this robust methodology to the 3 explanatory variables in the rat data for several values of the ellipsoidal coverage c. The results are confusing at best since they depend heavily on the chosen coverage c. At the recommended coverage $c = 11$, four observations are identified as outliers. Further it is only at $c = 12$ and 15 that case 2 is identified.

At this point we delete case 2. When we do so, the OLS regression of the response on the predictors is statistically null, as expected by the experimenter. The OLS residual plot, Figure 3, shows that the fitted values are generally between about .3 and .35, and the residuals between about -.1 and + .1, with no pattern. The LMS residual plot, Figure 4, gives a very different picture. Fitted values cover a much wider range, from .1 to .5, and the residuals are from about -.2 to +.2. The pattern of points is again troubling, but interpretation of it is not easy. Also, four points are identified as outliers, but this set does not include cases 0 and 18, found

Table 1: Explanatory variable outliers based on the minimum volume ellipsoid, rat data.

Coverage, c	Outliers
11	2, 4, 7, 12
12	2
13	4, 12, 15
14	4, 12
15	2
16	4, 12, 15
17	none
18	none

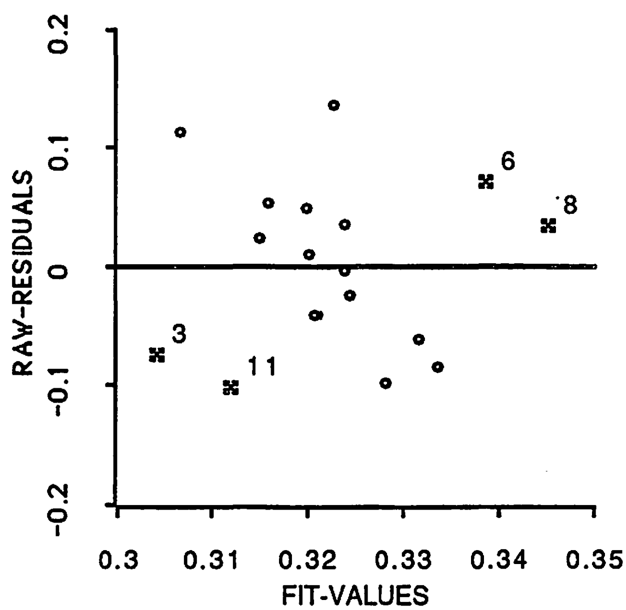


Figure 3: Rat data: OLS residuals versus fitted values, case 2 deleted. Points marked with an "×" are LMS outliers.

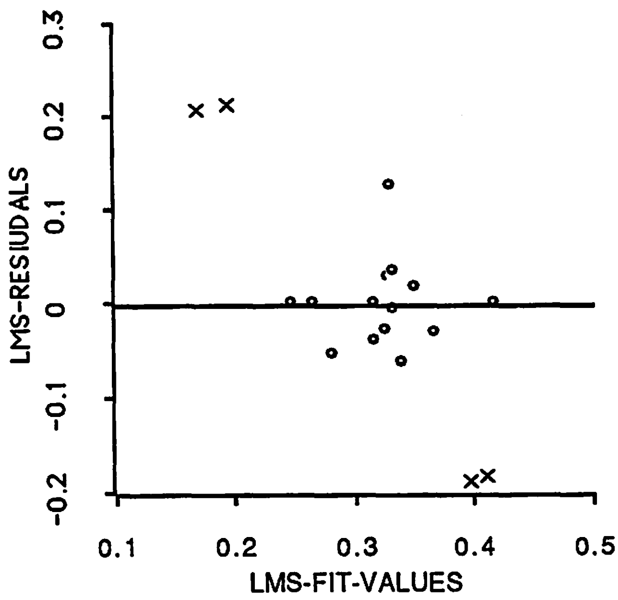


Figure 4: Rat data: LMS residuals versus fitted values, case 2 deleted. Points marked with an "×" are LMS outliers.

when case 2 was included. The LMS residuals are not equivalent to OLS residuals.

2.2 Lathe data

Reference: Weisberg (1985, p. 166). These data come from a central composite design with two factors, speed and feed. The response is log(tool life). The data have eight center points, 2 points at each of the corners, and one point at each star point. In a 3-D plot of {speed, log(life), feed}, curvature in the plot is evident, indicating that a first order linear model in speed and feed is not appropriate. One would want to be able to find such curvature by elementary residual plots.

Figure 5 is the OLS plot of residuals versus fitted values for the first order model. All the center points have negative residuals, and most other points have positive residuals. In combination with the curved trend in the plot, the need for some second order terms is apparent. Using LMS residuals, Figure 6, the diagnosis of curvature is far from clear. The center points all have residuals close to zero. Five of the other points (2 star points, 1 of a pair of corner points, and both of another

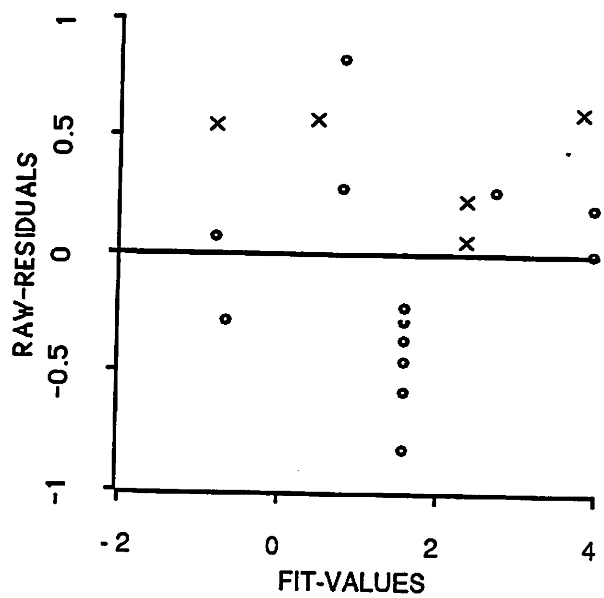


Figure 5: Lathe data: OLS residuals versus fitted values. Points marked with an "×" are LMS outliers.

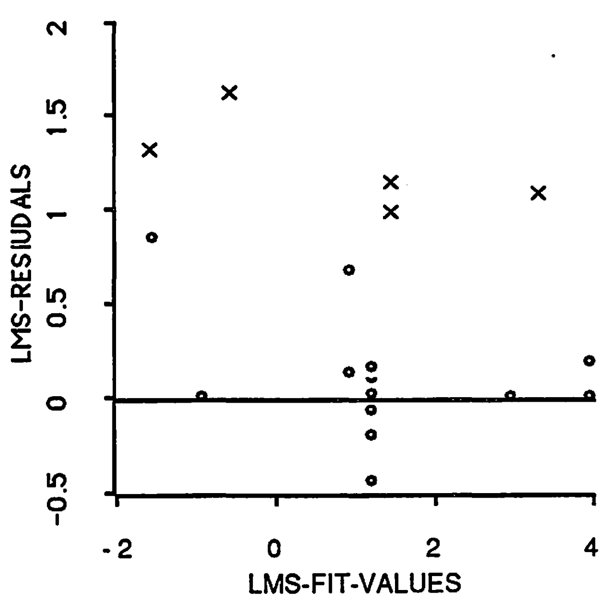


Figure 6: Lathe data: LMS residuals versus fitted values. Points marked with an "×" are LMS outliers.

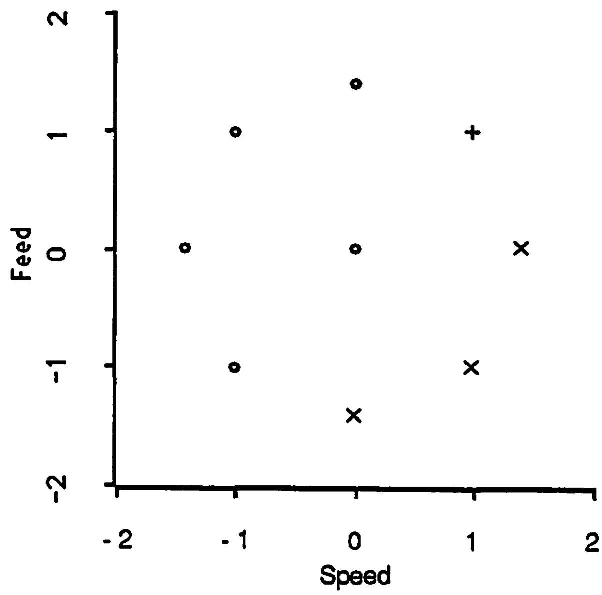


Figure 7: Lathe data: The design space, with LMS outliers marked

pair of corner points) are indicated as outliers.

The nature of the LMS fit is shown more clearly in Figure 7, which is a plot of the design points in the {speed, feed} plane. Points indicated by a '×' sign are one or more LMS outliers, while the point indicated by a '+' sign consists of one outlier and one point not an outlier. The LMS fit essentially ignores nearly half of the design space. In designs with replicated center points, the LMS fit will usually fit a plane through a representative value for the center points and a fraction of the design, ignoring much of the design space, or else the fit may simply ignore all the center points. One can force the fit to match the center points at the expense of fitting the design points by simply increasing the number of observations at the center point until it is nearly 50% of the total. In problems like these, the nature of the LMS fit, and the characteristics of the residuals, depend more on the design than they do on the model.

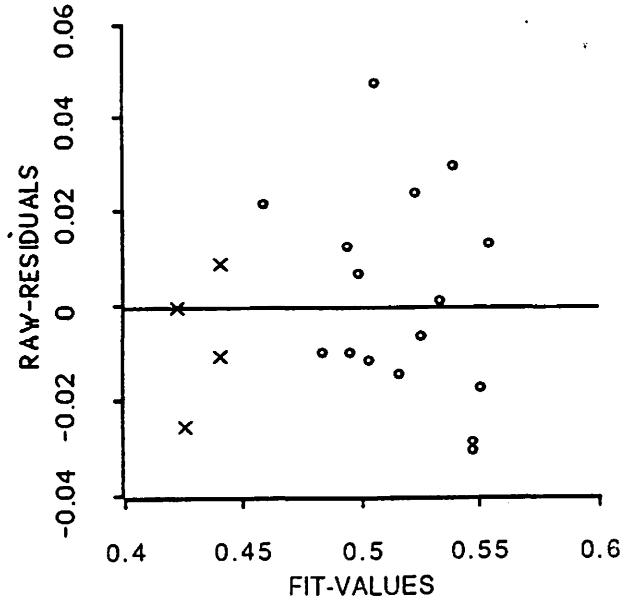


Figure 8: Wood gravity data: OLS residuals versus fitted values.

2.3 Modified wood gravity data

Reference: Rousseeuw and Leroy (1987, p. 243). Rousseeuw and Leroy (1987) modified the original wood gravity data given by Draper and Smith (1966, p. 227) by changing 4 of the x-vectors to form outliers in the factor space. These points are invisible in the standard OLS plot of residuals versus fitted values, Figure 8, but are visible in the LMS version in Figure 9. The marked points correspond in the two plots. It may be of some interest that the 4 suspect cases are at the extreme left in Figure 8 and they will be plainly evident for just about any 3-D rotation of OLS residuals. The OLS leverages (not shown) do not find this group of four points.

We next turn to detrended added variable plots (Cook and Weisberg, 1990), which are plot of residuals in which the x-axis is some other direction in the estimation space. Figure 10 is the OLS detrended added variable plot for the first and third predictors after the others (the direction chosen is then given by the change in fitted values for the model with all the predictors and the model excluding x_1 and x_3). The possible curvature in this plot may be indicative of a problem related to x_1 and x_3, such as the need to add a quadratic or an interaction to the model.

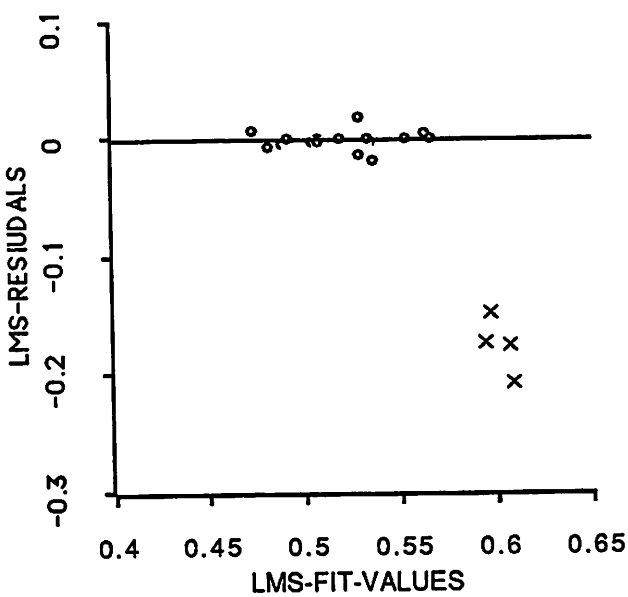


Figure 9: Wood Gravity data: LMS residuals versus fitted values.

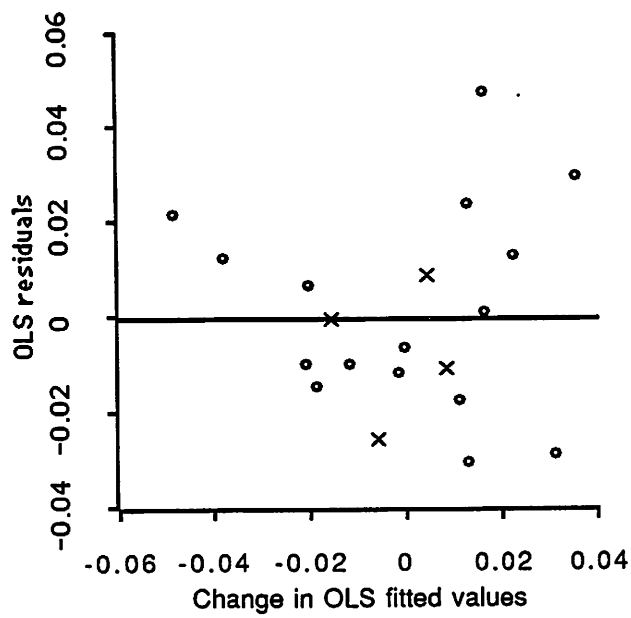


Figure 10: Wood gravity data: OLS detrended added variable plot for x_1 and x_3.

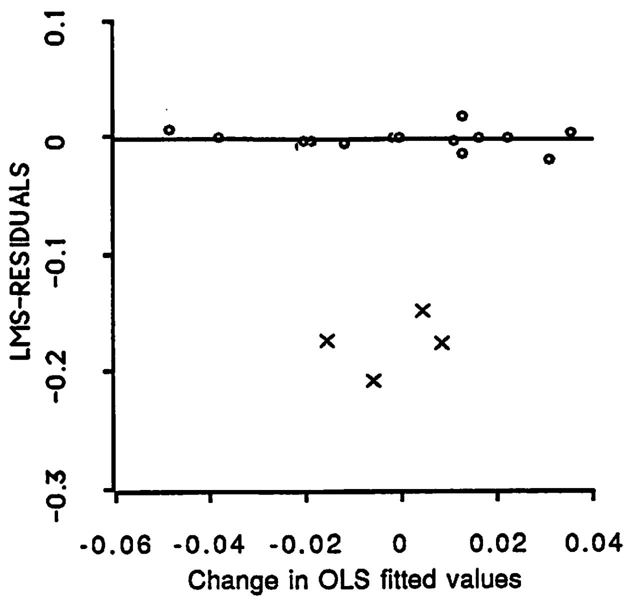


Figure 11: Wood gravity data: LMS residuals from the fit of all predictors versus the same direction as in the last figure.

Figure 11 is of the LMS residuals on the y-axis versus the same direction vector in the factor space as in Figure 10. The major feature of the LMS plot is again the four outliers, and these are so far removed from the remaining data that little else can be seen in the plot. To examine this plot further, it is useful to remove these four points, and redraw the plot to fill the frame with the remaining points. When this is done, we get Figure 12. Figure 12 shows more or less the same trend as does Figure 10 — possible curvature as a function of x_1 and x_3, although without first seeing Figure 10, one may well not view the trend in Figure 12 as sufficient for concern.

One is left with competing explanations for the apparent problems in these data: either we have a group of four outliers or we have an interaction. Both methods seem to be able to find either of these problems, with LMS more clearly finding outliers, and OLS more clearly indicating curvature. Either method may well lead to roughly the same conclusions if we are allowed to manipulate plots, and look at several different views of the data.

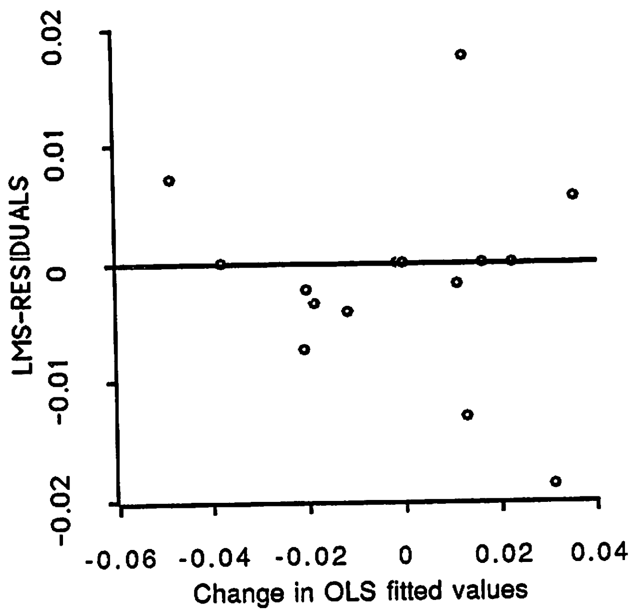


Figure 12: Wood gravity data: Same as last figure, but with the four outliers removed.

2.4 Cloud seeding data

Reference: Cook and Weisberg (1982), page 4. These data summarize the results of a cloud seeding experiment in Florida in the 1975. The predictors include several covariates that describe the day's weather and an action variable (0 if unseeded, 1 if seeded). The response is log(rainfall) in a target area. The unit of analysis is a day. Twenty-four days are included in the data. Our previous analysis leads to the following conclusions:

1. Case 1 is a highly influential case, as it is quite different from the rest of the data.

2. Cases 6 and possibly 23 are outlying; both are unseeded days with too little observed rainfall.

3. Most importantly, an action by covariate interaction seems important. The crucial conclusion is that the effect of seeding depends on the covariates.

Table 2: Explanatory variable outliers based on the minimum volume ellipsoid, cloud seeding data.

Coverage, c	Outliers
11	1, 6, 22
12	1, 2, 5, 8, 22
13	1, 5, 8
14	0, 1, 3, 5, 6, 22, 23
15	1, 2, 5, 8, 15, 17, 22, 23
16	0, 1, 5, 6, 17
17	0, 1, 5, 17
18	0, 1, 5, 17
19	0, 1, 5, 17
20	0, 1, 5, 17
21	1
22	1
23	1
24	14

Case 1 is identified as influential by using standard least squares deletion diagnostics: Cook's distance (see, for example, Cook and Weisberg, 1982) $D_1 = 1.88$, while the second largest distance is $D_6 = .79$. For LMS, case 1 is in the elemental set used to define the estimator, so the residual for this case is zero. As in the rat data of the first example, we computed the MVE to get a "robust" measure of leverage. Case 1 is identified as outlying when the number of points included in the MVE is $c = [(24 + 6 + 1)/2] = 15$, or for virtually any value of c. As shown in Table 2, at $c = 15$, seven additional cases are identified as outliers, while for other values of c the number and location of the outliers is very different. It is unclear to us how to interpret an analysis that may declare fully one-third of the data to be outlying.

Of course, leverage is not the same as influence. Rousseeuw and van Zomeren (1990) suggest that a measure of influence can be obtained graphically by plotting the robust leverage versus the LMS residual; cases that are large on both are influential, while cases with large leverage but small residuals are called "good" leverage

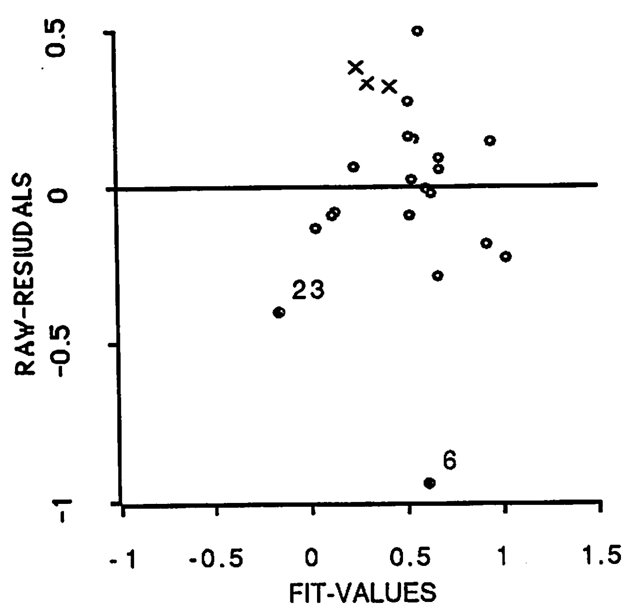


Figure 13: Cloud seeding data: OLS residuals versus fitted values.

points, since they appear to match the trend of the data. Thus, case 1 would be called a "good" leverage point, since it is in the basis that determines the LMS estimator. This is certainly at odds with the OLS diagnostics result that labels case 1 as influential; indeed, if case 1 is deleted from the data and the model is refit, both the LMS and OLS regressions show substantial changes.

Figure 13 gives the standard plot of residuals versus fitted values for the fit of the response on all the predictors with all the data included, and Figure 14 gives the equivalent plot based on LMS regression. As is usual, LMS gives a much wider range both to the fitted values and to the residuals. Both identify case 6 as being well separated from the bulk of the data; however, LMS also declares cases 7, 12 and 15 to be outliers. Interestingly, the other "canonical" outlier, case 23, is in the elemental subset that is used to fit the LMS regression, so it has a residual of zero. The OLS residuals in Figure 13 do not suggest much special about 7, 12, and 15, although the shape of this plot may be suggestive of curvature.

The two methods seem to give a different view on outliers, LMS finding more. We now turn to the more important question: finding the need for an action by

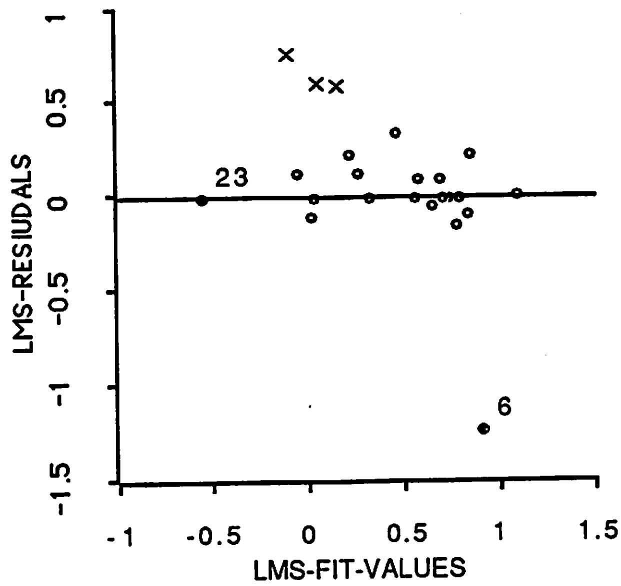


Figure 14: Cloud seeding data: LMS residuals versus fitted values.

covariate interaction. Cook and Weisberg (1989) use a three dimensional residual plot to find the interaction. The plot is of x-axis = (difference between fitted values for fitting the model and fitting the model without the action variable); z-axis = (fitted values from full model without action); and y-axis = (OLS residuals from full model). An action by covariate interaction will be revealed by a bowl or saddle shape in this plot. Figure 15 gives one two-dimensional projection of this three-dimensional plot showing the required shape. Figure 16 shows the LMS residuals on the y-axis, with the other two axes exactly the same as in Figure 15. While the outlying case 6 is in evidence in this plot, any hint of the interaction is lost.

We view the x- and z-axes as providing directions in the p-dimensional estimation space, so as we spin we view residuals projected onto various directions. This is relevant regardless of the estimation method. One might argue here that the equivalent 3-D plot for LMS would use LMS fitted values to define the other axes. This may well be justifiable, but the theory for doing this does not exist.

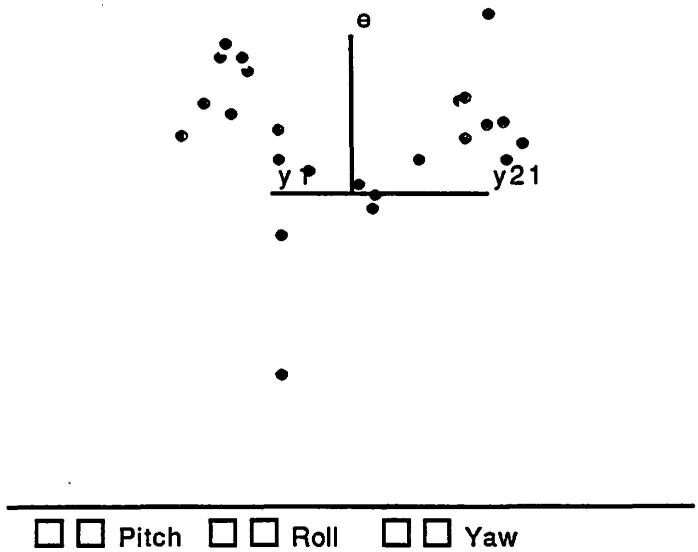


Figure 15: Cloud seeding data: OLS 3-D plot for *action*

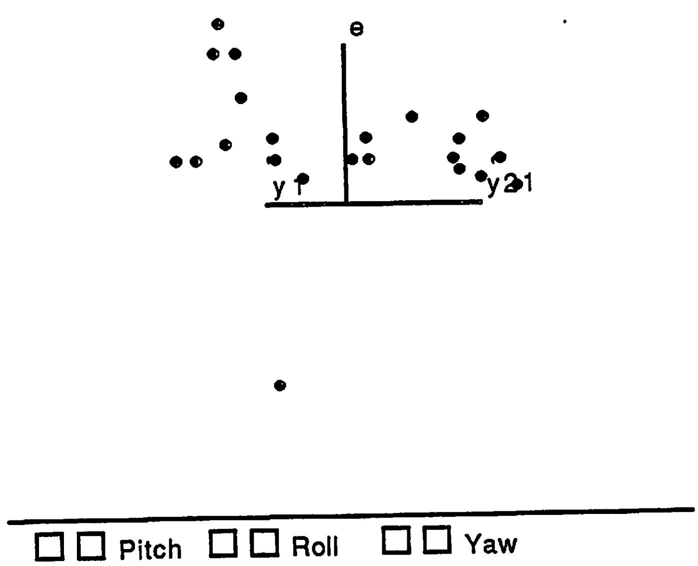


Figure 16: Cloud seeding data: LMS 3-D plot for *action*

Table 3: PVC data

x	y	x	y	x	y	x	y
796.33	0	72.77	6.552	956.93	146.89	715.50	162.90
139.69	2.26	1174.24	529.20	254.73	16.65	166.11	42.34
248.22	43.97	449.57	8.76	258.05	10.06	211.54	2.92
909.94	176.76	46.17	0	109.83	7.63	314.82	1.01
1268.85	155.17	373.28	11.84	159.82	13.74	473.72	13.70
423.53	14.39	204.75	0	84.46	6.18	327.75	14.08
164.29	23.61	3184.03	0.63	544.84	98.16	66.28	0
86.81	0.53	337.50	100.36	65.25	4.03	111.54	10.20
565.75	19.54	67.48	49.33	1133.94	0	1008.98	73.85
48.45	0.49	593.28	0	798.14	11.79	665.36	7.83
339.01	13.92	192.19	157.69	88.68	0.12	1447.47	96.54
341.15	3.39	1122.60	1442.70	750.57	172.07	235.77	16.59
62.40	8.89	1464.02	53.55	878.35	87.63	129.36	6.12
162.53	19.44	168.14	45.17	365.46	1.47	174.92	2.99
1517.96	187.69	480.42	82.72	92.79	0.10	720.31	0
574.90	3.90	549.02	24.51				

2.5 Premature Ventricular Contractions

A sample of 62 patients were all treated with an identical drug. Recorded for each patient was x = scaled baseline rate of premature ventricular contractions (PVC's) and the response is y = scaled PVC's at three months post treatment. The experiment included a control group, but it is not reported here. The data are given in Table 3.

The goal is to obtain a predictive model of y as a function of x. As with most problems with a single predictor, the plot of x *versus* y is highly informative, and is shown as Figure 17. From this plot, it is unclear if any useful predictive model is possible, since no strong trend is apparent. However, both x and y take values over several orders of magnitude, suggesting that these quantities should be transformed, a suggestion we will persue shortly. First, however, we shall explore fitting OLS and LMS to these untransformed data.

Figures 18 and 19 are the plots of fitted values *versus* residuals for OLS and LMS

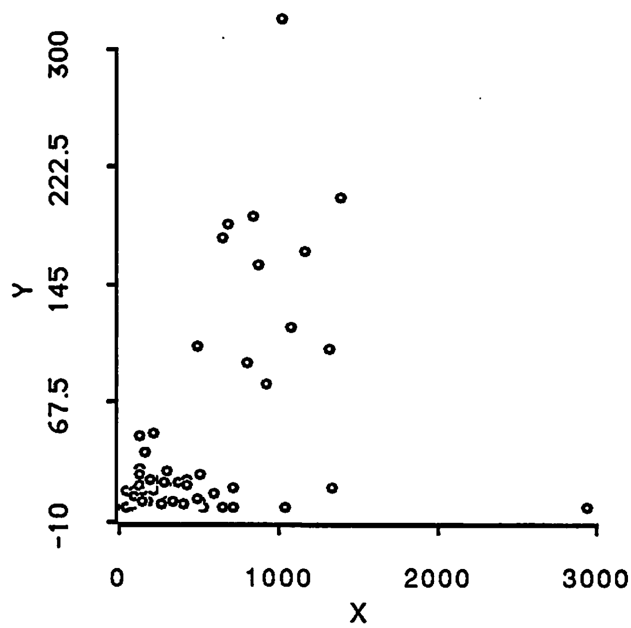


Figure 17: PVC data: Plot of the raw data.

respectively. These two plots are strikingly different, because the OLS and LMS fits are very different. The OLS fit has intercept 8.0 and slope of 0.053, thereby attempting to match the data while paying attention to the points with relatively large values of y. Using standard one-at-a-time outlier methodology (e.g., Weisberg, 1985, Chapter 5), the cases marked with an "×" and a "□" are labelled as likely outliers. However, the main feature of this plot is its wedge shape. This shape can generally be ascribed to one of two causes. First, it could be indicative of heteroscedasticity, with variance increasing as a function of y. The second cause is somewhat more subtle, but nonetheless common in practice. Suppose that the response y has some fixed maximum or minimum value; for example, y may be bounded below by zero. If there are many ponts in the data with differing x values for which the bound is achieved, or nearly achieved, then a wedge shape will result. The extreme example of this occurs when the response takes on only the values 0 and 1, so the plot of fitted values *versus* residuals is just two lines. In this example, the lower bound in the plot can probably be attributed to the fixed boundary at zero, while the upper trend may be indicative of heteroscedasticity. Both of

Figure 18: PVC data: OLS fitted-values *versus* residuals.

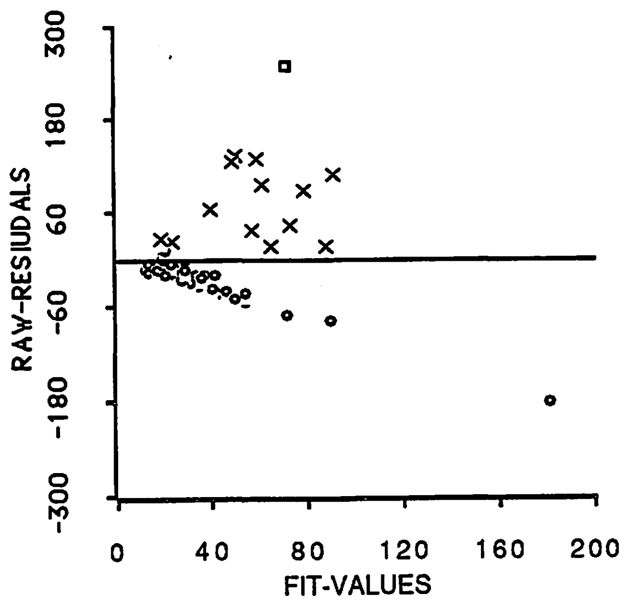
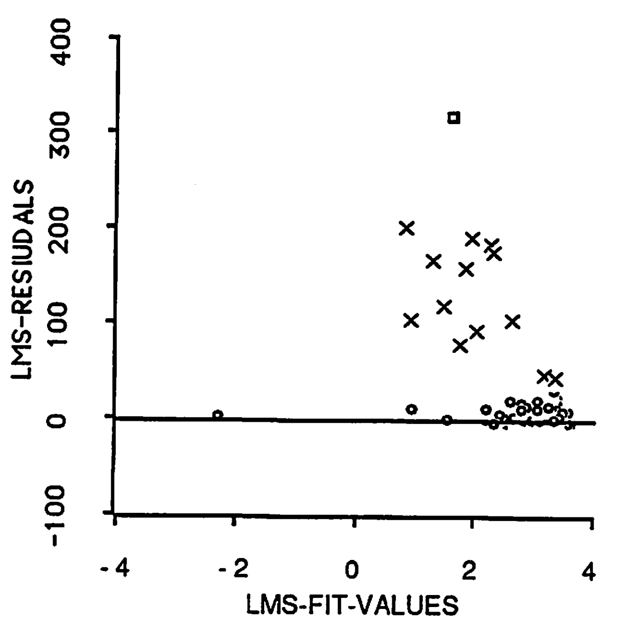


Figure 19: PVC data: LMS fitted-values *versus* residuals.

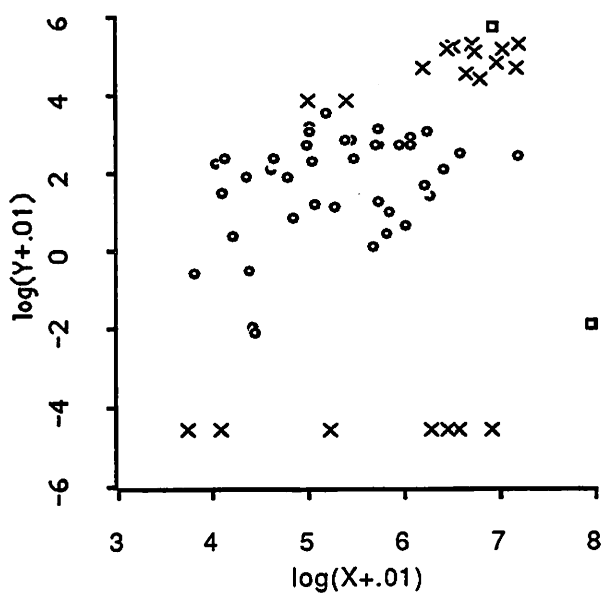


these problems can be handled by transformations. For example, a log transform of $y + c$ will expand the lower end of the scale and contract the upper, thereby simultaneously handling both problems (c is a small constant required because $y = 0$ for some cases).

The LMS plot in Figure 19 is quite different. The fitted slope and intercept are 4.0 and -0.0016, respectively. Thus LMS essentially fits the points with y near zero, since they make up more than half the data, while ignoring the rest. Points in Figure 19 marked by a "+" or a "□" are LMS outliers; only the case marked by the "□" is both an LMS and OLS outlier. Unlike the OLS fit, the LMS fit will depend on the relative proportion of cases with y close to zero. If that proportion had been somewhat smaller, the fit would have been different, and the shape in Figure 16 would probably be different. Figure 19 may in fact contain useful information about the need to transform to get a more sensible model, but one cannot use OLS intuition applied to LMS to infer this.

Figure 20 is a plot of $\log(x)$ *versus* $\log(y + 0.01)$. The horizontal band of points at the bottom of the plot are the points for which $y = 0$; by changing c, these points can be moved up or down at will. It is apparent that, apart from these seven cases, and possibly the case marked with a "□" on the plot, there is a reasonably strong relatonship between $\log(x)$ and $\log(y)$. Both LMS and OLS label these eight cases as outliers; both methods give essentially the same fitted model. We would conclude that (1) for some cases $y = 0$ regardless of the value of x, assuming that the data are a suitably chosen random sample from a target population; (2) for the majority of cases, $\log(x)$ and $\log(y)$ are linearly related; and (3) one outlier is apparent, which had a small, nonzero value of y corresponding to the largest value of x.

Figure 20: PVC data: log plot.



3 Conclusions

We like to view least squares fitting of linear regression models as a general diagnostic procedure, largely because it is sensitive to virtually all the assumptions made in a regression problem. For example, least squares residuals and related statistics can be used to find outliers, curvature, nonconstant variance, and other modelling problems. LMS, on the other hand, achieves high breakdown, but at a very high cost. As in the lathe data, for example, LMS can fit a model that is apparently wrong, and thereby miss the most important feature of these data: the curvature. In accommodating outlying points, the importance of other assumptions in modelling is distorted.

In working these examples, we in fact used four estimators: OLS and LMS, as discussed throughout this paper, the Huber M-estimate with $c = 1.345$, and the default robust estimator computed by the *New S* package. For these examples, the latter two gave results that were essentially identical to OLS. This is only to be expected, since these robust estimators are "least squares in the middle", and none of the examples, except perhaps the cloud seeding data, have very large residuals. However, one can also find examples that would make these estimates give results that are difficult to use and interpret (some results are given by Cook and Weisberg, 1982, Section 5.5).

The examples that we presented were selected to emphasize the potential differences between OLS and LMS diagnostic methodology. Not all data sets result in such diametrically opposed conclusions.

We would be reluctant to advise the general use of standard diagnostic methodology when applied to robust estimation procedures until, and unless, adequate theory for this is devised. Even the idea of first doing a robust fit to find the outliers, and then doing least squares on the rest, may not be useful general advice. In the lathe data, for example, if this were done then the curvature would never be found. Of course, one is not limited to a single estimation method: robust methods in concert with OLS may well lead to very powerful analyses.

4 References

Cook, R. D. and Weisberg, S. (1982). *Residuals and Influence in Regression.* London: Chapman-Hall.

Cook, R. D. and Weisberg, S. (1989). "Regression diagnostics with dynamic graphics (with discussion)", *Technometrics, 31*, 277-311.

Cook, R. D. and Weisberg, S. (1990). "Added variable plots in linear regression", in Stahel, W. and Weisberg, S., eds., *IMA Proceedings on Robustness and Diagnostics.* New York: Springer, in press.

Draper, N. and Smith, H. (1966), *Applied Regression Analysis.* New York: Wiley.

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

Krasker, W. S. and Welsch, R. E. (1982). "Efficient bounded-influence regression estimation." *Journal of the American Statistical Association, 77*, 595-604.

Rousseeuw, P. J. and Leroy, A. (1987). *Robust Regression and Outlier Detection.* New York: Wiley.

Rousseeuw, P. J. and van Zomeren,B. C. (1990). Unmasking multivariate outliers and leverage points. *Journal of the American Statistical Association*, to appear.

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