

Not to be cited without
permission of the authors¹

Canadian Atlantic Fisheries
Scientific Advisory Committee

CAFSAC Research Document 83/58

Ne pas citer sans
autorisation des auteurs¹

Comité scientifique consultatif des
pêches canadiennes dans l'Atlantique

CSCPCA Document de recherche 83/58

Diagnosis for the multiplicative model

by

G.N. White
Marine Fish Division
Bedford Institute of Oceanography
Fisheries and Oceans
P.O. Box 1006
Dartmouth, N.S. B2Y 4A2

and

S. Gavaris
Northwest Atlantic Fisheries Center
Fisheries and Oceans
P.O. Box 5667
St. John's, Newfoundland A1C 5X1

¹ This series documents the scientific basis for fisheries management advice in Atlantic Canada. As such, it addresses the issues of the day in the time frames required and the Research Documents it contains are not intended as definitive statements on the subjects addressed but rather as progress reports on ongoing investigations.

Research Documents are produced in the official language in which they are provided to the Secretariat by the author.

¹ Cette série documente les bases scientifiques des conseils de gestion des pêches sur la côte atlantique du Canada. Comme telle, elle couvre les problèmes actuels selon les échéanciers voulus et les Documents de recherche qu'elle contient ne doivent pas être considérés comme des énoncés finals sur les sujets traités mais plutôt comme des rapports d'étape sur les études en cours.

Les Documents de recherche sont publiés dans la langue officielle utilisée par les auteurs dans le manuscrit envoyé au secrétariat.

Abstract

The problems of diagnosis for the multiplicative model are shown to be primarily ones of leverage and influence. Partial regression leverage plots are suggested as an appropriate diagnostic tool, however, the costs of producing these plots is large and existing data flows do not allow scientists to examine the original data for possible problems. A method which reduces the cost of partial regression leverage plots while providing improved access to the data is proposed.

Résumé

Les principaux problèmes diagnostiques posés par le modèle multiplicatif en sont de point-levier et d'influence. On estime que des courbes de régressions partielles des points-leviers seraient un outil de diagnostic approprié. Cependant leur coût de production est élevé et les systèmes actuels de traitement des données ne permettent pas aux scientifiques d'examiner si les données originales peuvent poser des problèmes. On propose une méthode qui réduit le coût de production de ces courbes tout en améliorant l'accès aux données.

Introduction

Indices of abundance are fundamental elements for the scientific management of fisheries. While there are many classes of data that can be used to develop abundance indices, commercial fisheries are often important sources of information. Catch and effort data from commercial fisheries often provide longer time series and more extensive coverage than can be obtained by other means. On the other hand, data from commercial fisheries are subject to a variety of biases and inadequacies over which scientists have little control.

Commercial catch and effort data are becoming increasingly important for the activities of the CAFSAC Groundfish Subcommittee. Abundance indices from research vessel surveys have exhibited patterns of increasing variability with increased stock abundance (unpublished data). Research vessel surveys in Atlantic Canada have also suffered from the inevitable problems of changes in equipment (Koeller and Smith, 1983). Requests for information and analyses relating to the activities of the Task Force on Atlantic Fisheries have included such items as projections of commercial catch rates. Although clearly a key element in attempts to rationalize the industry, such projections require estimates of catch rates from all major segments of the fishery. Thus they are not compatible with the common practice of developing abundance indices from data for a restricted set of gears, areas, and months.

The practice of employing restricted data sets can be criticized on the grounds that it ignores much of the available data, introduces elements of subjectivity, and can be unduly influenced by the effects of changes in fishing patterns or technology. The multiplicative model (Robson, 1966; Gavaris, 1980), was developed with the goal of allowing information from a variety of categories to be incorporated into abundance indices. It also provides estimates of relative fishing power that might be usefully employed in economic investigations.

The multiplicative model, while no doubt an oversimplification of the true relationships between various categories of effort, is capable of explaining a substantial fraction of the variation in many sets of catch-effort data. Thus, while the multiplicative model may someday be replaced by more realistic models such as those discussed by Laurec (1977), it can be expected to play a substantial role in fisheries management for some time to come.

The practical use of the multiplicative model presents significant problems. Its design is dictated more by the way fisheries data is collected and processed than by the dictates of good or at least conventional statistical practice. In particular, the multiplicative model employs a regression consisting entirely of dummy variables. This design, while precluding some of the problems that

occur in using conventional regression models, creates other problems. The multiplicative model can result in many of the parameters being determined primarily by a small number of data points. This occurs when the total number of categories is large in relation to the number of observations. In statistics, this is known as the problem of "influential data", a topic that has not yet entered into the discussions provided by most introductory textbooks on regression.

Problems of influential data in regression are easily diagnosed using graphical techniques when there are only two carriers (this paper employs the terminology of Mosteller and Tukey, 1977). For regression problems with a large number of carriers the partial regression leverage plot (Larsen and McLeary, 1972) or component and component plus residuals plot (Daniel and Wood, 1980) provides a useful diagnostic tool, although the computational effort is high if diagnostic plots are created for all the carriers. The development of criteria to pinpoint carriers for diagnostic examination is currently the subject of active research.

This paper has two goals. The first is to demonstrate that collinearity is not a practical concern in the use of the multiplicative model. This has important consequences for the choice of numerical procedures used to solve the regression problem. The second is to give a preliminary report of our findings regarding the problem of leverage. Our central conclusions are:

1. leverage problems can be expected when there are few observations in which a particular carrier is involved;
2. existing software available to assessment biologists (SPSS, SAS, BMDP) does not provide the means to obtain diagnostics currently considered necessary to deal with leverage problem (some diagnostics can be obtained using MINITAB); and
3. partial regression leverage plots can be a useful diagnostic tool, although we are unable to offer definitive criteria for the selection of carriers to be examined.

Collinearity in the Multiplicative Model

The problem of collinearity does not arise (except as a trivial case) until one considers regressions with more than two carriers. Consider, for example, the regression model

$$y = b_1x_1 + b_2x_2 + \epsilon \quad (1)$$

where the carriers are x_1 and x_2 , and ϵ is the associated stochastic effect. This model can be visualized in the 3-dimensional (x_1, x_2, y) space as a plane (Figure 1). If the data

$$\{(x_{i1}, x_{i2}, y_i)\}_{i=1}^n$$

consist entirely of observations for which x_{i1} and x_{i2} lie close to a line, as do the first ten points in Figure 1, then the plane defined by (1) is poorly determined. The consequences of colinearity include both a loss of numerical precision in carrying out the calculations and also large statistical uncertainties associated with the estimates for some of the parameters.

It is useful at this stage to observe that a single "maverick" observation such as point number 11 in Figure 1 would eliminate the colinearity. Such an observation should, however, be subject to careful examination. Maverick observations often arise as a consequence of errors in processing data. Furthermore, a small change in such an observation will greatly alter the fitted regression surface by causing it to shift around the line determined by the remaining observations. This effect is known, for obvious reasons, as "leverage", and cannot be detected from conventional analyses of regression residuals because the residual at an influential point will be small. The problem of leverage will be discussed further below.

Returning to the problem of colinearity, it is important to note that the design of the multiplicative model requires the carriers to be vectors of 0's and 1's. Consider, for example, the situation illustrated in Table 2, which is patterned after the example in Robson's (1966) original description of the multiplicative model. If the first gear category and year are chosen as standards for comparison, then $p_1 = 0$ and $q_1 = 0$, leaving a total of five parameters to be determined from the regression model $\underline{y} = X\underline{b}$, where \underline{b} is defined as

$$\underline{b} = (C, p_2, p_3, q_2, q_3)$$

and X has five columns and a row for each observation.

If all of the nine possible observations were available, then, taking the cells of Table 2 from left to right in successive rows:

$$X = \begin{matrix} & 1 & 0 & 0 & 0 & 0 \\ & 1 & 0 & 0 & 1 & 0 \\ & 1 & 0 & 0 & 0 & 1 \\ & 1 & 1 & 0 & 0 & 0 \\ & 1 & 1 & 0 & 1 & 0 \\ & 1 & 1 & 0 & 0 & 1 \\ & 1 & 0 & 1 & 0 & 0 \\ & 1 & 0 & 1 & 1 & 0 \\ & 1 & 0 & 1 & 0 & 1 \end{matrix} \quad (2)$$

The use of the two constraints, $p_1 = q_1 = 0$ ensures that the columns of X are linearly independent, provided there are observations in all cells. Frequently, however, not all cells have data. Consider, for example, the case where gear three was observed only in year 3, and the other gears were not observed in year 3. Then

$$X = \begin{matrix} & 1 & 0 & 0 & 0 & 0 \\ & 1 & 0 & 0 & 1 & 0 \\ X = & 1 & 1 & 0 & 0 & 0 \\ & 1 & 1 & 0 & 1 & 0 \\ & 1 & 0 & 1 & 0 & 1 \end{matrix} \quad (3)$$

has two identical columns (3 and 5) corresponding to the parameters p_3 and q_3 associated with gear 3 and year 3. These data cannot provide unique estimates for both p_3 and q_3 . Even if one parameter is omitted so that the regression problem can be solved, it is impossible to say how much of the effect of the remaining parameter is due to the gear and how much is due to, for instance, a change in abundance of the stock being fished. The matrix in (3) demonstrates precise colinearity of columns 3 and 5. Although some linear regression programs will cheerfully report a solution to such problems (generally these programs estimate $b = p_3 + q_3$ and report the values of p_3 and q_3 such that $b = p_3 + q_3$ and $p_3 + q_3$ is minimized), most abort after issuing an error message.

The colinearity problem which results when some observations do not overlap with others is of concern primarily for the problems of interpretation that arise when a single parameter combines several different effects. Because of the special nature of the design matrix, X , in the multiplicative model, the problem of near colinearity that can occur with more general models does not arise. Thus it is not necessary to use regression software designed to handle problems with near collinearity. The only caveat to this statement is that rounding errors could conceivably make an exactly colinear matrix appear acceptable to some regression programs. Thus it is desirable to detect cases of exact colinearity before attempting to carry out a regression. Since this occurs when there is no overlap between gear categories in area or time, it is generally easy to detect given a knowledge of the fishery.

Leverage and Influence

Leverage, as discussed in the previous section, refers to the potential effect on regression estimates of maverick observations. In this section we will concentrate on a graphical method for investigating the effect of such observations on estimates for regression parameters. The principle tool for such investigations is the partial regression leverage plot. Because each partial regression leverage plot requires an additional regression calculation, it is

certainly desirable to develop criteria that can be used to pinpoint observations whose leverage is dangerously high. Recent literature on the problem of leverage diagnostics has been largely concerned with the search for such criteria, but no entirely satisfactory results are currently available.

The idea of using partial regression leverage plots is relatively new. The earliest published description of these plots is that of Larsen and McCleary (1972). Daniel and Wood (1980) introduced a variant which includes the predicted line, calling them "component and component plus residual plots". A partial regression leverage plot is based on the comparison of two sets of residuals R_j and $R(j)$, where R_j is the vector of partial X residuals obtained by regressing the j -th column of X on the remaining columns. The partial y residuals, $R(j)$, is the vector of residuals obtained by omitting column j from the original regression for y . The latter vector can be calculated from the full regression as

$$R(j) = R + b_j X_j$$

where R is the vector of residuals from the full regression and X_j is the j -th column of X . It can be shown that the slope of a least-squares regression line through the origin on the plot of $R(j)$ against R_j will be b_j (Mosteller and Tukey, 1977). Thus one can easily visualize the effect of any single observation on the estimate for b_j as is done for single variable regressions. In this regard it is very helpful to follow Daniel and Wood's (1980) strategy and include the regression line in the plot.

Partial regression leverage plots for the data in Table 1 are shown in Figures 2-4. In each case, the maverick observation clearly plays the determining role in the estimates for the parameters. It should be noted that, because the regression line passes close to the maverick point, this determining role cannot be detected using conventional residual plots.

The application of partial regression leverage plots to problems with many parameters would be greatly simplified if some criteria were available to select the observations with the greatest leverage for examination. A number of criteria have been proposed in the literature. These invariably involve the diagonal elements of the "hat" matrix, H , which is defined formally as the $n \times n$ matrix:

$$H = X(X^T X)^{-1} X^T$$

This matrix is too large to be computed in its entirety for problems involving more than a couple hundred observations. While it is possible to calculate the individual diagonal elements of H , few statistical packages are designed to provide this information (Hoaglin and Welsch, 1978).

In order to get some feeling for the behaviour of the hat matrix diagonal elements, simulated catch-standardization data were generated for $n = 200$, $p = 30$. Leverage is associated with large values in the diagonal of H , but there is currently no accepted criterion for levels that would be too high. The average value of the diagonal elements is p/n , but the level $2p/n$ suggested by Belsley et al. (1980) is considered by others to be too high (Huber, 1983; Krasker and Welsch, 1982). In the simulations it was found that the largest values of h_{ii} were generally, but not always, associated with the columns of X that had the fewest 1's. This is intuitively quite reasonable since, if few observations affect a given parameter, each will play an important role. It should, however, be noted that the number of observations does not entirely explain the leverage values, so that it is highly desirable to calculate the diagonal elements of H .

In the absence of quantitative criteria to serve as a guide in detecting influential observations, we have relied heavily on partial regression leverage plots. In particular we have examined an actual data set consisting of 305 observations in which one observation (no. 287) was found to have been coded as 58 (where " " indicates a blank) in the effort field. This coding may be read variously as 58 or 580 depending on the language, compiler, and operating system being used. Table 3 presents the parameter estimates obtained for the two possible values of effort in record no. 287. It can be seen that two parameter estimates (variables 10 and 38, corresponding to $B[11]$ and $B[39]$) change by almost one standard error. It is clearly important to have diagnostic procedures which can draw our attention to any observation that can have such a major effect on the results. Because observation no. 287 has high leverage, it does not stand out (indeed, it is lost in the crowd) in conventional residual plots (Figures 5, 6). It does, however, stand out in certain of the partial regression leverage plots (Figures 7, 8). Note, however, that other observations also appear to exert considerable influence. Of particular concern is the partial regression leverage plot for variable 38 ($B[39]$) representing the final year. This parameter is directly related to the abundance estimate that is most critical to tuning a sequential population analysis.

Here we face a dilemma. Having identified several observations which appear to depart from the trend in the rest of the data, we would like to examine them in more detail. In most cases, we will not find so obvious an ambiguity as was the case for observation number 287 in our example.

In summary, then, parameter estimates obtained using the multiplicative model, as with other forms of regression analysis are sensitive to maverick observations. While partial regression leverage plots appear to be useful in identifying these observations, they do not tell us what to do about these observations. Such decisions should

be based on an investigation of the underlying data which contribute to a particular aggregated regression datum. This has led us to propose a novel approach to regression diagnosis in the following section.

It is important to note that partial regression leverage plots do not show the weight placed on a particular observation. Changes in the weights used for a regression can dramatically affect the influence of an observation with high leverage. This effect can be judged by comparing partial regression leverage plots (with regression line included) for both weighted and unweighted regressions.

A New Approach to Diagnostics for the Multiplicative Model

The multiplicative model presents a rather special type of regression problem. The principle problems that must be guarded against when using this model are related to leverage and influence. In the previous section it was noted that the current inability to examine the components of the individual regression datum makes it difficult to deal with the problem of leverage.

Leverage and influence fall outside the domain of existing statistical packages. The software that has been developed for investigating such problems is of recent design and thus assumes that the user has access to large central memory or virtual memory adequate to hold data and working arrays. The most straight forward approach to developing practical solution methods for the multiplicative model would be to use the raw (disaggregated) data. This, however, requires virtual memory and, in any case, represents a wasteful expenditure of computing resources. For the multiplicative model, the estimates of the regression parameters will not be altered if suitable estimates of location for the individual observations in each category are used in a much smaller (properly weighted) regression. The real value of including all the data is for diagnostic analyses. This purpose can be met using the regression coefficients obtained from the aggregated data to calculate residuals for the disaggregated data.

The procedure described in the preceeding paragraph considerably reduces the computational effort required to obtain residuals or partial residuals. Thus it becomes computationally feasible to examine a large number of partial regression leverage plots. This fact reduces the need for accurate selection criteria, thus reducing one of the key difficulties in the analysis of leverage problems. Indeed, from our experience the perspective gained from examining a large number of partial regressions can be quite helpful.

This approach raises a number of issues that require further study:

1. Is the current method of aggregation (i.e. simply summing values of catch and effort) appropriate when the data are to be used with the multiplicative model?
2. Would it be useful to weight the regression by the number of observations contributing to each regression datum?
3. Are existing graphical displays adequate? In particular, how can a single observation be identified from a plot?

Acknowledgement

The authors are grateful to M. Mingo for assistance in preparing this report.

References

- Belsley, D.A., E. Kuh, and R.E. Welsch. 1980. Regression diagnostics: identifying influential data and sources of collinearity. Wiley, N.Y.
- Daniel, F.C. and F.S. Wood. 1980. Fitting equations to data. 2nd Edition. Wiley, N.Y.
- Gavaris, S. 1980. Use of a multiplicative model to estimate catch rate and effort from commercial data. Can. J. Fish. Aquat. Sci. 37: 2272-2275.
- Hoaglin, D.C. and R.E. Welsch. 1978. The hat matrix in regression and ANOVA. Amer. Stat. 32: 17-22.
- Huber, P.J. 1983. Minimax aspects of bounded-influence regression. J. Am. Stat. Assoc. 78: 66-79.
- Koeller, P. and S.J. Smith. 1983. Preliminary analysis of A.T. Cameron - Lady Hammond comparative fishing experiments 1979-1981. In preparation.
- Krasker, W.S. and R.E. Welsch. 1982. Efficient bounded-influence regression estimation. J. Am. Stat. Assoc. 77: 595-604.
- Larsen, W.A. and S.J. McCleary. 1972. The use of partial residual plots in regression analysis. Technometrics 14: 781-790.
- Laurec, A. 1977. Analyse et estimation des puissances de peche. J. Cons. Int. Explor. Mer 37: 173-185.
- Mosteller, F. and J.W. Tukey. 1977. Data analysis and regression - a second course in statistics. Addison-Wesley, Reading, Mass.
- Robson, D.S. 1966. Estimation of the relative fishing power of individual ships. ICNAF Res. Bull. 3: 5-14.

Table 1. Artificial data set for Figures 1-4.

Observation No.	1.00	2.00	3.00	4.00	5.00	6.00	7.00	8.00	9.00	10.00	11.00
X1	0.90	1.00	1.90	2.00	2.90	3.00	3.90	4.00	4.90	5.00	5.00
X2	1.00	1.00	2.00	2.00	3.00	3.00	4.00	4.00	5.00	5.00	0.00
Y	0.68	0.70	1.35	1.40	2.08	2.10	2.78	2.80	3.48	3.50	YEND

PARAMETER ESTIMATES:

Y 11	[B 1]	[B 2]
0	0.00024	0.70
1	0.20	0.50
2	0.40	0.30

Table 2. Multiplicative model for the catch rates of three gears in three years. Based on the example of Robson (1966). Entries in table are the model terms for the natural log of catch rate.

Gear	Year		
	1	2	3
1	$c + p_1 + q_1 + e_{11}$	$c + p_1 + q_2 + e_{12}$	$c + p_1 + q_3 + e_{13}$
2	$c + p_2 + q_1 + e_{21}$	$c + p_2 + q_2 + e_{22}$	$c + p_2 + q_3 + e_{23}$
3	$c + p_3 + q_1 + e_{31}$	$c + p_3 + q_2 + e_{32}$	$c + p_3 + q_3 + e_{33}$

p_i = effect of gear

q_i = effect of year

e_{ij} = stochastic error

Table 3a. Regression coefficients and analysis of variance for multiplicative model. a. EFFORT [287]=58.

REGRESSION OF MULTIPLICATIVE MODEL

MULTIPLE R,.....0.650
MULTIPLE R SQUARED,....0.422

ANALYSIS OF VARIANCE

SOURCE OF VARIATION	DF	SUMS OF SQUARES	MEAN SQUARES	F-VALUE
INTERCEPT	1	1.504E1	1.504E1	
REGRESSION	38	6.586E1	1.733E0	5.109
TYPE 1	9	3.326E1	3.742E0	11.090
TYPE 2	11	1.479E1	1.344E0	3.963
TYPE 3	8	4.501E0	5.626E-1	1.659
TYPE 4	10	1.003E1	1.003E0	2.957
RESIDUALS	266	9.022E1	3.392E-1	
TOTAL	305	1.711E2		

REGRESSION COEFFICIENTS

VARIABLE	COEFFICIENT	STD. ERROR	NO. OBS.
INTERCEPT	-0.201	0.231	305
1	-0.080	0.160	19
2	-0.145	0.138	27
3	-0.020	0.168	18
4	0.367	0.119	59
5	-0.583	0.177	18
6	-0.771	0.290	5
7	0.680	0.116	57
8	0.534	0.272	6
9	0.937	0.188	17
10	-0.118	0.250	13
11	-0.160	0.227	24
12	-0.192	0.233	19
13	-0.019	0.220	29
14	0.252	0.227	26
15	0.146	0.215	40
16	0.231	0.217	37
17	0.123	0.218	34
18	-0.060	0.215	39
19	-0.162	0.239	18
20	-0.062	0.242	16
21	-0.128	0.182	17
22	0.073	0.176	19
23	-0.377	0.172	20
24	-0.292	0.167	23
25	-0.426	0.204	12
26	-0.399	0.276	6
27	-0.517	0.254	7
28	0.231	0.255	7
29	-0.596	0.213	11
30	-0.418	0.226	10
31	-0.141	0.198	18
32	-0.434	0.187	18
33	-0.170	0.229	11
34	-0.765	0.211	12
35	-0.664	0.242	8
36	-0.216	0.172	25
37	-0.274	0.166	27
38	-0.120	0.179	20

Table 3b. Regression coefficients and analysis of variance for multiplicative model.
b. EFFORT [287]=580.

REGRESSION OF MULTIPLICATIVE MODEL

MULTIPLE R,.....0.638

MULTIPLE R SQUARED,....0.406

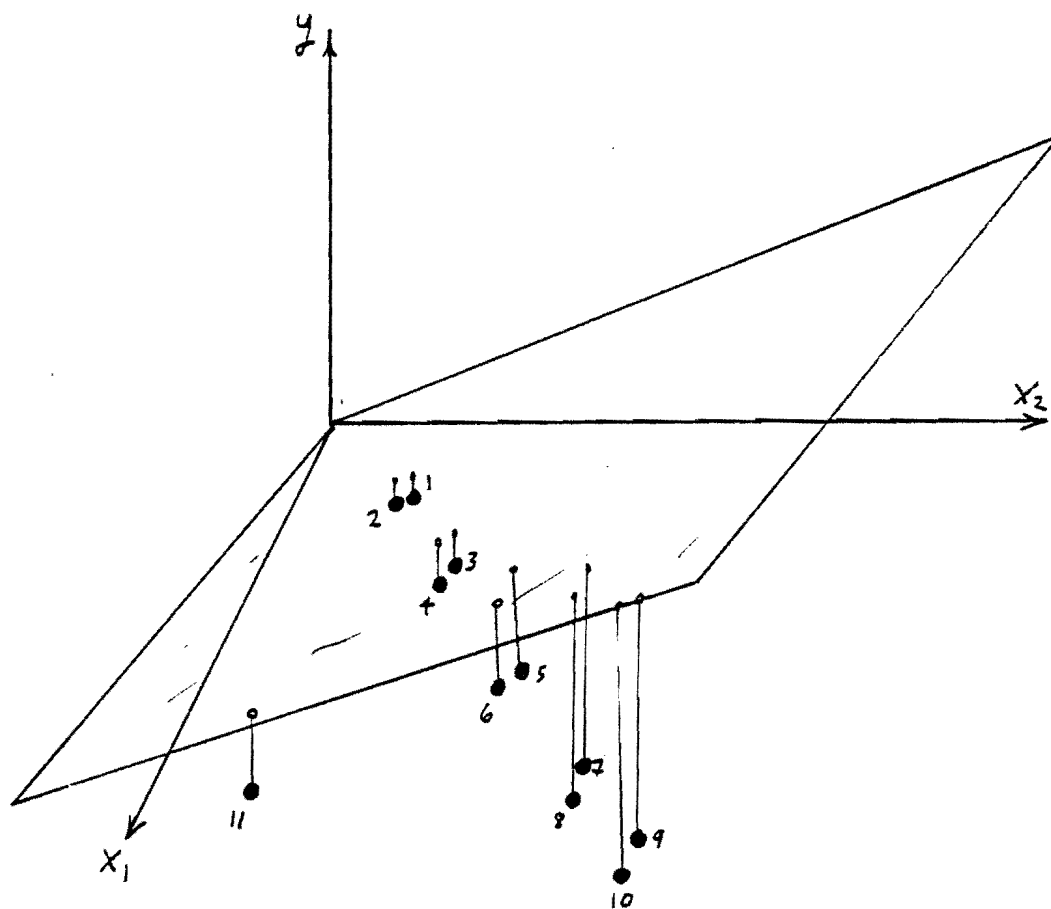
ANALYSIS OF VARIANCE

SOURCE OF VARIATION	DF	SUMS OF SQUARES	MEAN SQUARES	F-VALUE
INTERCEPT	1	1.608E1	1.608E1	
REGRESSION	38	6.442E1	1.695E0	4.794
TYPE 1	9	3.226E1	3.585E0	10.138
TYPE 2	11	1.521E1	1.383E0	3.910
TYPE 3	8	4.661E0	5.826E-1	1.648
TYPE 4	10	1.014E1	1.014E0	2.867
RESIDUALS	266	9.406E1	3.536E-1	
TOTAL	305	1.745E2		

REGRESSION COEFFICIENTS

VARIABLE	COEFFICIENT	STD. ERROR	NO. OBS.
INTERCEPT	-0.160	0.236	305
1	-0.086	0.163	19
2	-0.154	0.140	27
3	-0.012	0.171	18
4	0.358	0.122	59
5	-0.608	0.181	18
6	-0.803	0.296	5
7	0.633	0.118	57
8	0.513	0.277	6
9	0.923	0.192	17
10	-0.303	0.255	13
11	-0.166	0.231	24
12	-0.196	0.238	19
13	-0.025	0.224	29
14	0.241	0.232	26
15	0.135	0.219	40
16	0.219	0.222	37
17	0.107	0.222	34
18	-0.075	0.220	39
19	-0.185	0.244	18
20	0.051	0.247	16
21	-0.137	0.186	17
22	0.064	0.179	19
23	-0.400	0.175	20
24	-0.294	0.171	23
25	-0.420	0.208	12
26	-0.416	0.282	6
27	-0.541	0.259	7
28	0.205	0.260	7
29	-0.601	0.217	11
30	-0.404	0.230	10
31	-0.144	0.202	18
32	-0.439	0.191	18
33	-0.171	0.234	11
34	-0.780	0.216	12
35	-0.679	0.247	8
36	-0.223	0.176	25
37	-0.293	0.170	27
38	-0.237	0.183	20

Figure 1. Three dimensional graph of the regression model (1).
Observation no. 11 is a 'maverick' point with a large
leverage value. Data are given in Table 1.



- ⁿ - observation in the (x_1, x_2) plane
- ⊥ - location in the plane determined by the regression model $y = b_1x_1 + b_2x_2$

Figure 2. Partial regression leverage plots for data of Table 1
with $Y[11]=0$.

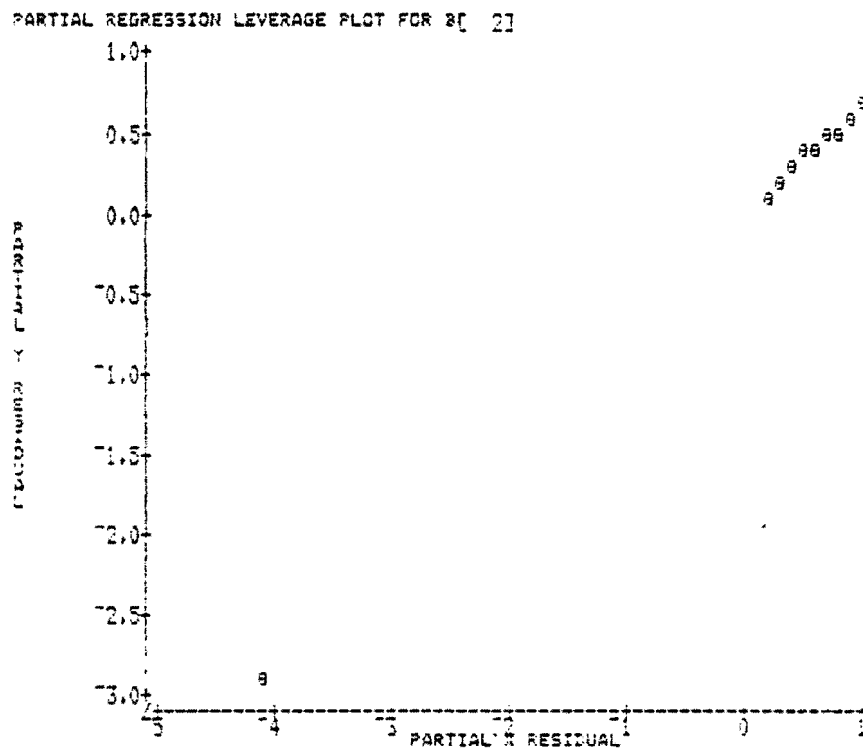
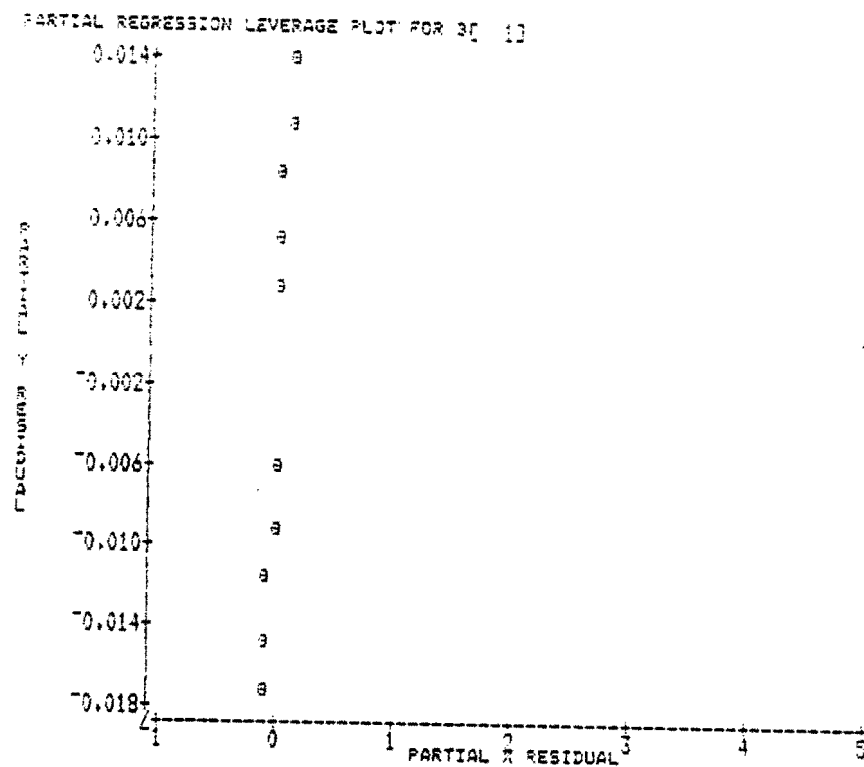


Figure 3. Partial regression leverage plots for data of Table 1 with $Y[11]=1$.

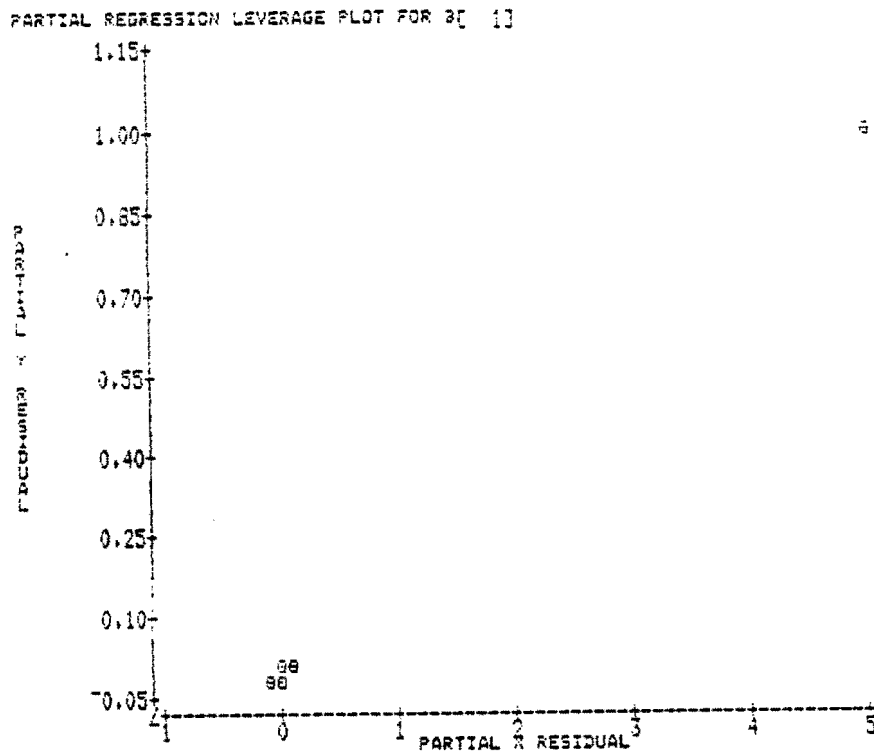


Figure 4. Partial regression leverage plots for data of Table 1 with $Y[11]=2$.

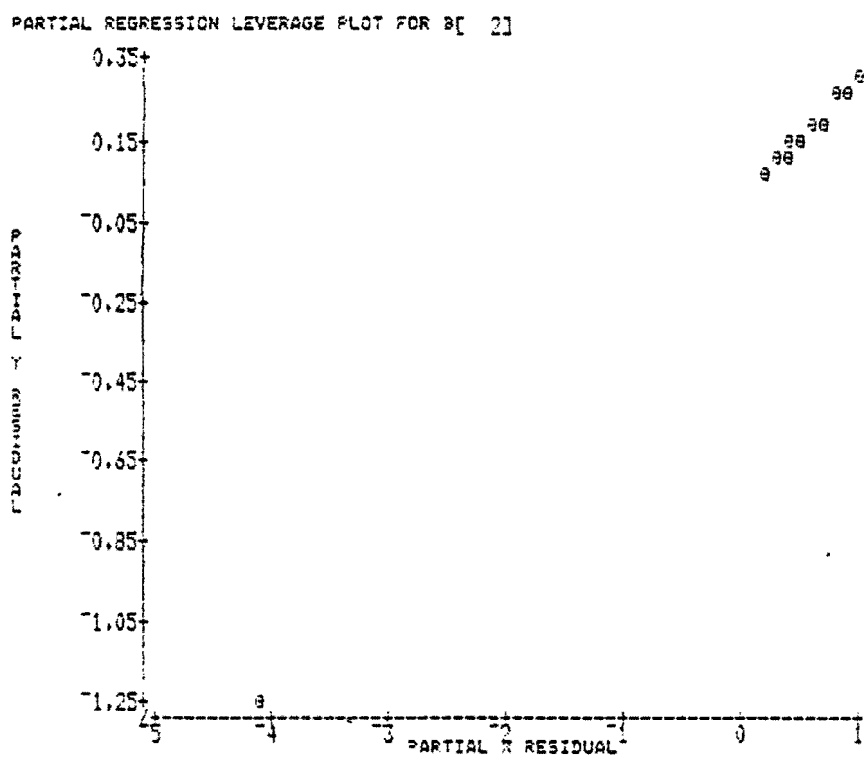
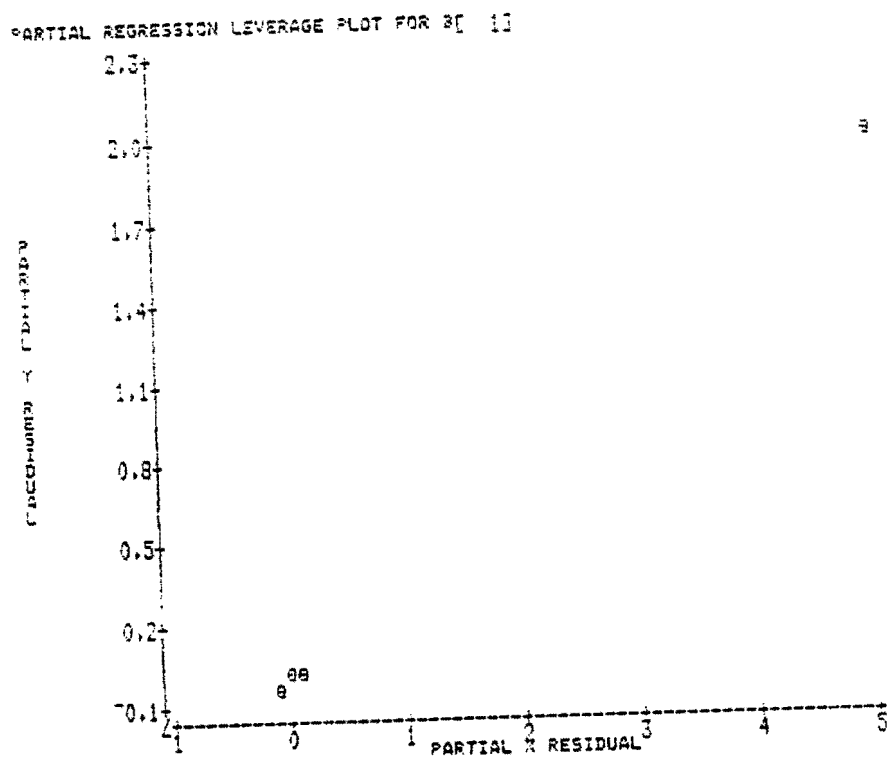


Figure 5a. Expected normal value plots for regressions of Table 3a.
a. EFFORT [287]=58.

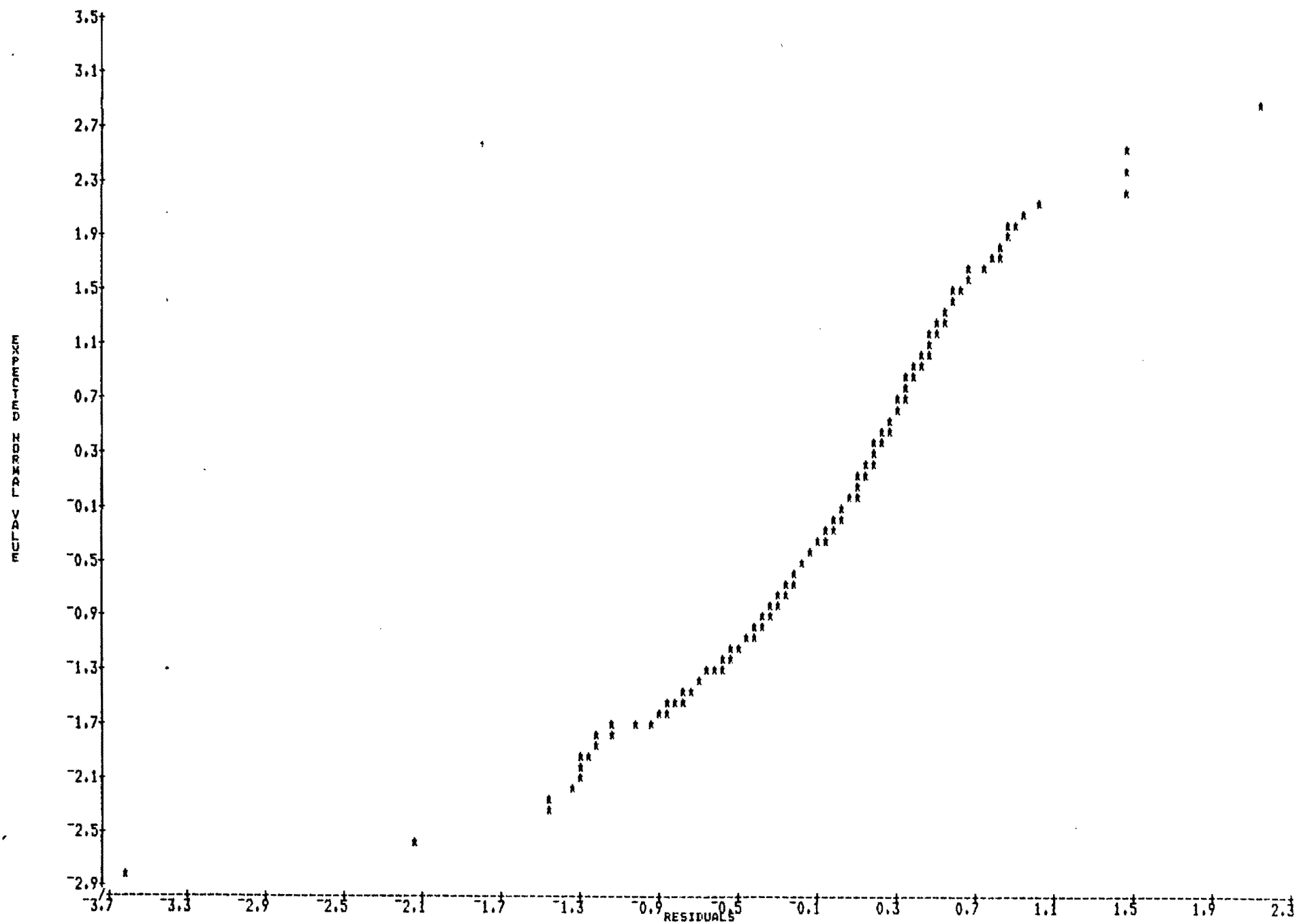


Figure 5 b. Expected normal value plots for regressions of Table 3b.
 b. EFFORT [287]=580

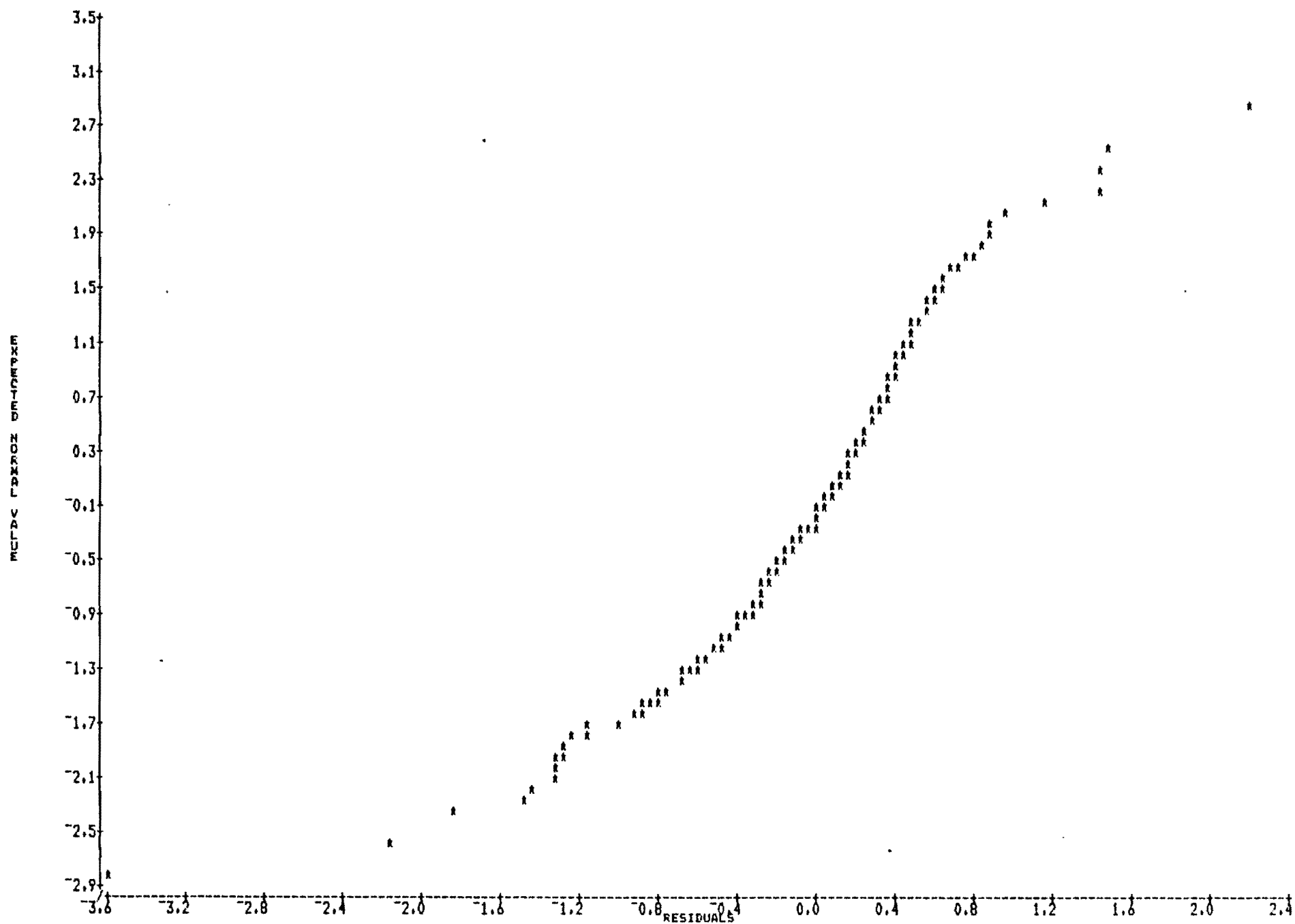


Figure 6a. Residuals vs. predicted in catch plots for regressions of Table 3a.
a. EFFORT [287]=58.

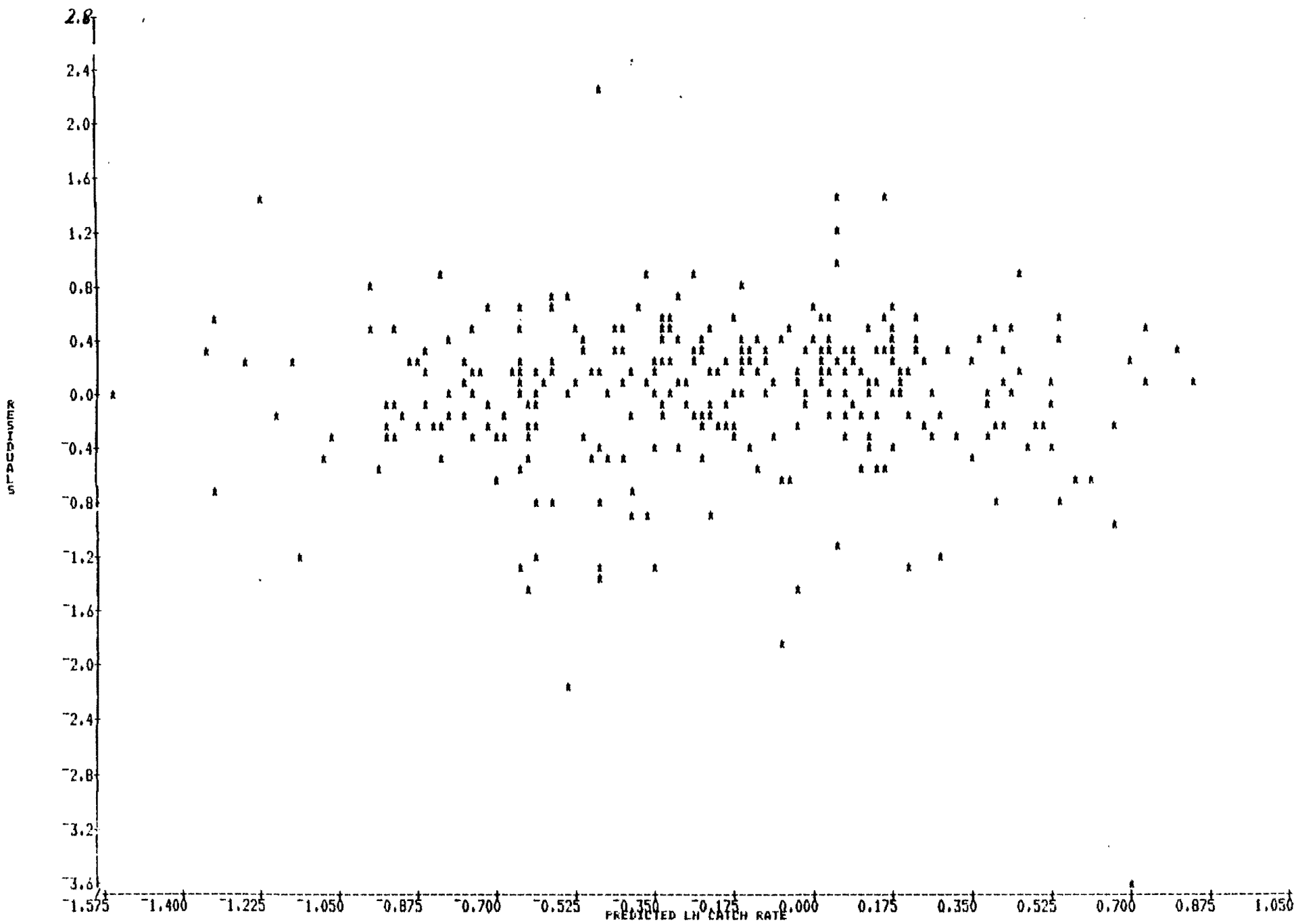


Figure 6b. Residuals vs. predicted in catch rate plots for regressions of Table 3b.
b. EFFORT [287]=580.

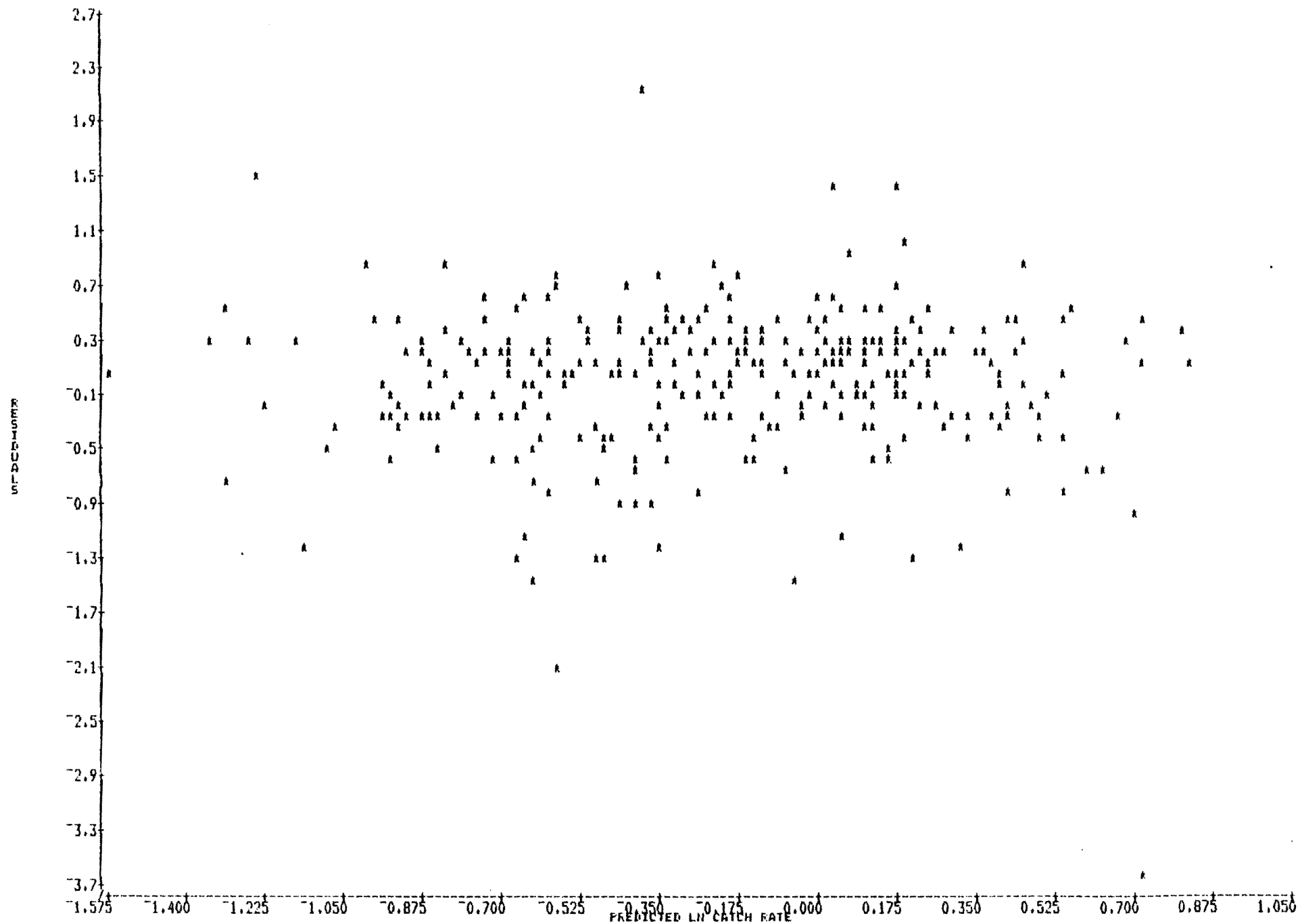


Figure 7a. Partial regression leverage plots for variable 10 (B[11]) in regressions of Table 3a.

a. EFFORT [287]=58.

PARTIAL REGRESSION LEVERAGE PLOT FOR B[11]

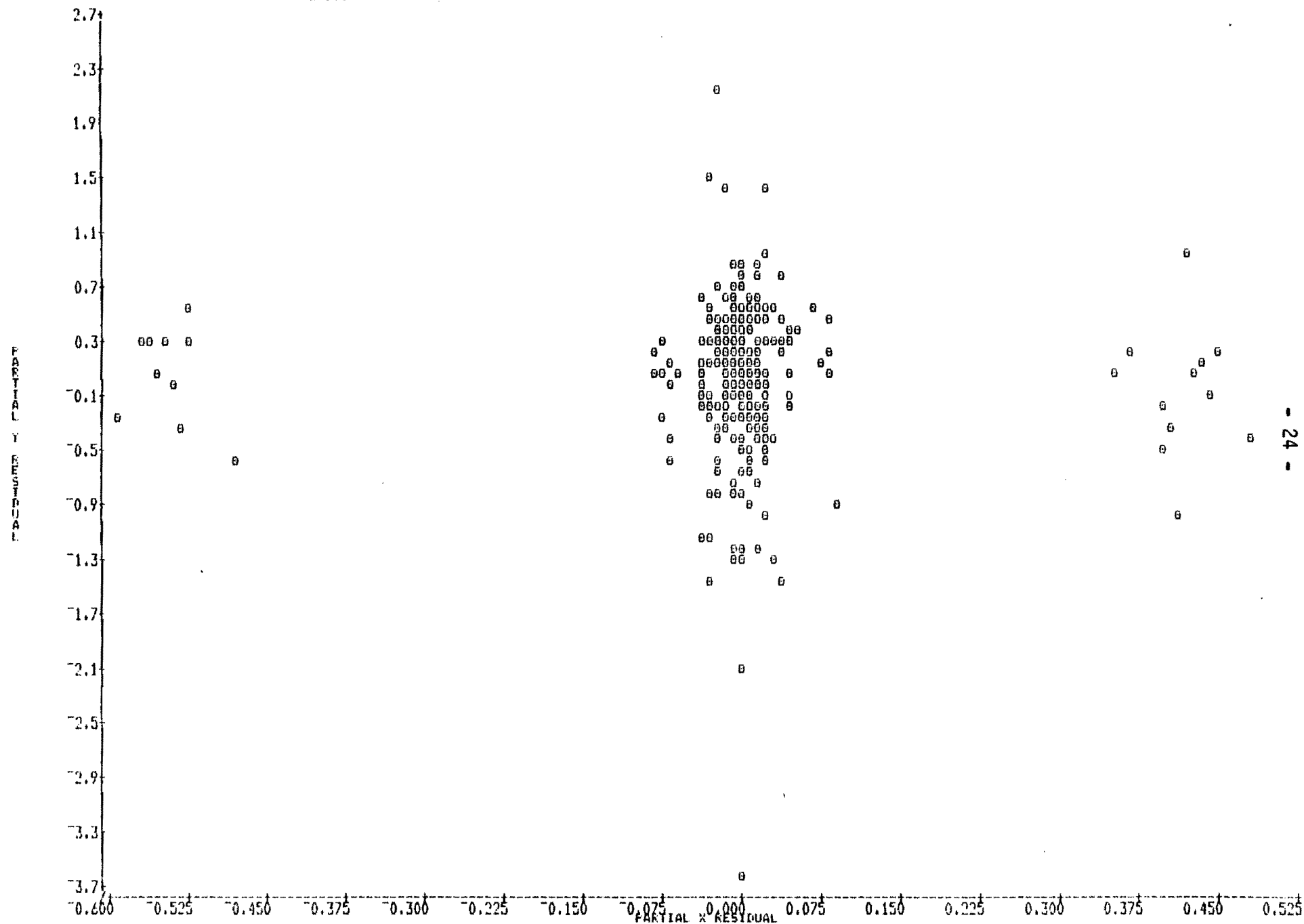


Figure 7b. Partial regression leverage plots for variable 10 (B[11]) in regressions of Table 3b.

h. EFFORT [287]=580.

PARTIAL REGRESSION LEVERAGE PLOT FOR B[11]

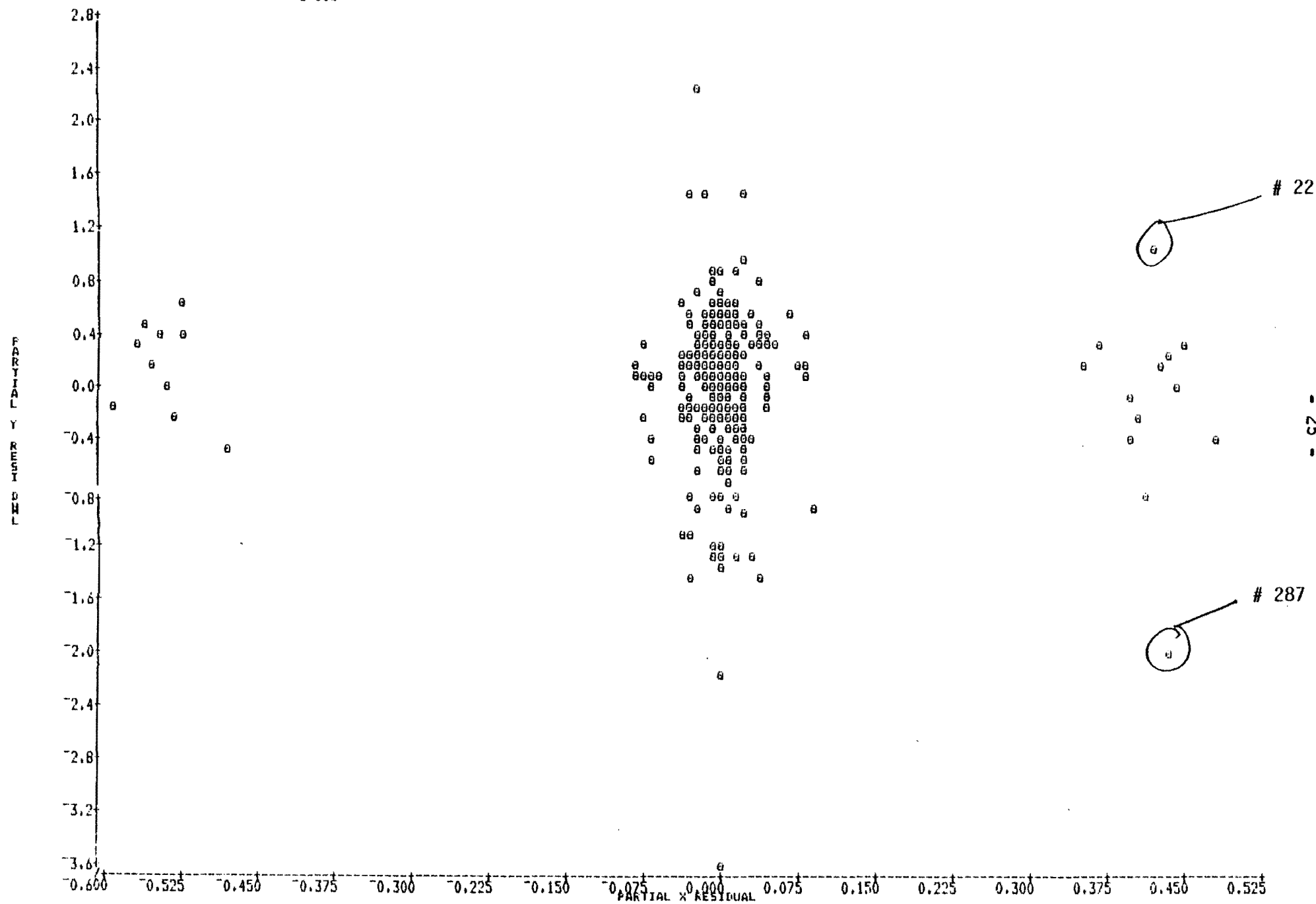


Figure 8a. Partial regression leverage plots for variable 38 (B[39]) in regressions of Table 3a.
a EFFORT [287]=58.

PARTIAL REGRESSION LEVERAGE PLOT FOR B[39]

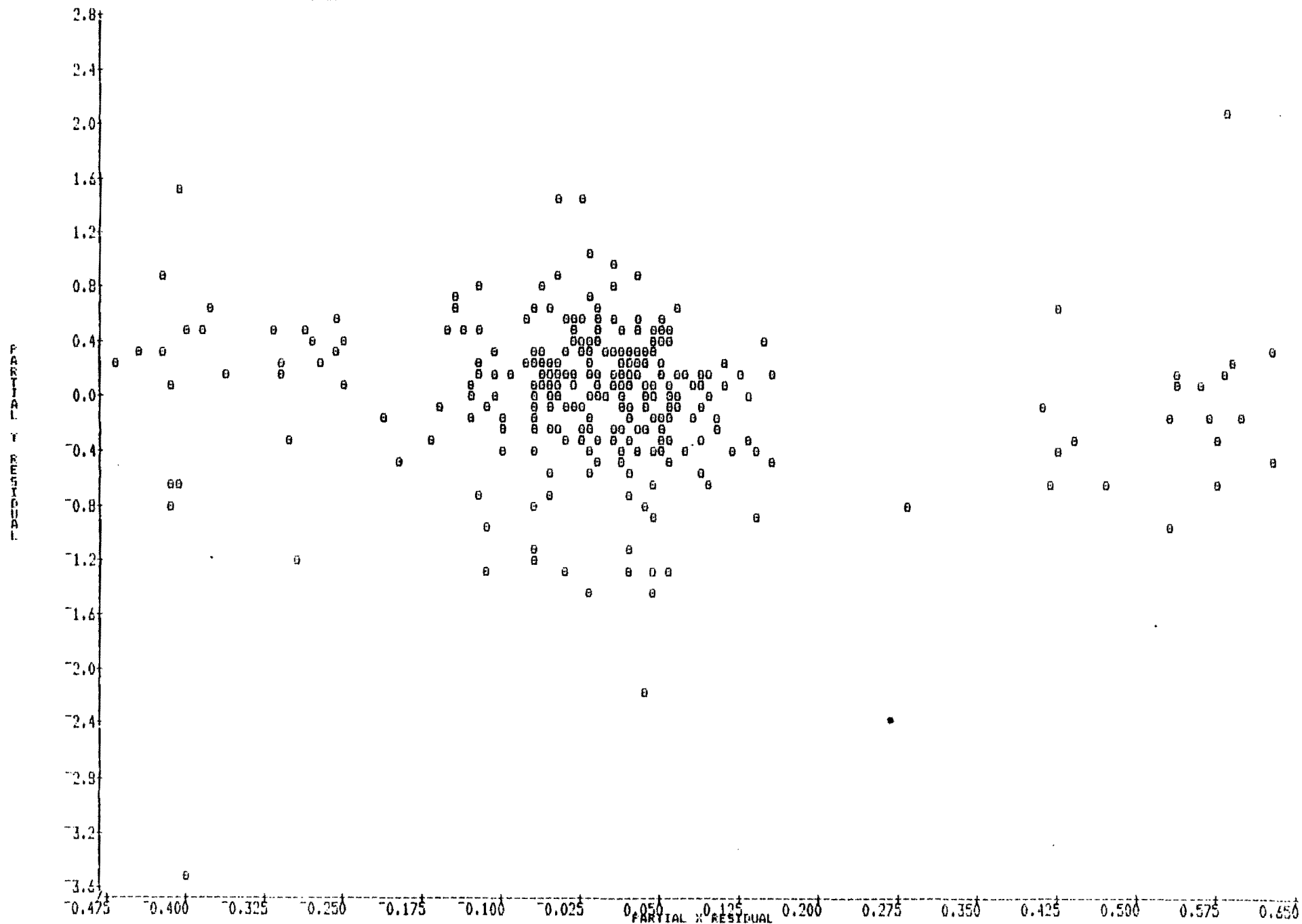


Figure 8b. Partial regression leverage plots for variable 38 (B[39]) in regressions of Table 3b.
b. EFFORT [287]=580.

