

**USING
MEASUREMENT ERROR MODEL
TO STUDY THE SHELTER EFFECT**

by

TONY Y. S. LEUNG

Master of Science project

April 1992

Department of Statistics and Applied Probability

University of Alberta

Edmonton, Alberta

Canada

ACKNOWLEDGEMENT

I am grateful to my supervisor, Dr. D. Wiens, for his suggestions and comments throughout the entire project. I would like to thank Dr. J. Wilson for his advice on the meteorological aspect of this project, and Mr. T. Thompson for helping me in collecting the data. I also like to thank Dr. R. Zamar for his helpful suggestions on orthogonal regression estimation. I also like to thank all the graduate students in the department for their support.

Finally, I am indebted to my parents for their support and encouragement in the past twenty five years.

1. Introduction

The development of vegetation during various stages of growth depends on several meteorological conditions. One of the conditions is the winds; strong and persistent winds will prevent vegetation growth because they reduce soil moisture and physically damage the vegetation. In areas where winds have such hazards, brushes of trees can be installed to produce a shelterbelt in order to reduce wind speeds. Many findings have proven that the areas near the shelterbelt and opposite the wind will have the shelter effect, i.e. higher temperature.

In Ellerslie weather station, a brush of evergreen trees is planted northward from the instrument shelter, see figure 1.1. We want to find out whether the trees have a shelter effect on the temperature at that station. In order to study the effect, we have chosen the Edmonton International Airport weather station to help our study. Since two stations are only about 10 miles apart, we expect that readings from both stations should be very close. Ten years' daily records of mean daily temperature, total hours of bright sunshine, total km of winds run from the north for both stations are available. Using all ten years observations, we use the standard test concerning the difference between two population means to test whether the mean temperature at Ellerslie is significantly higher than the mean temperature at the airport station. This test and the corresponding estimates have concluded that the temperature at Ellerslie is on average higher than the temperature at the airport. We also run the test using the days with total north wind greater than 10 km; the same conclusion is drawn. Detailed discussion with a meteorologist has led to a conjecture that the shelter effect may be the cause of the higher temperature at Ellerslie. Under this conjecture, we expect that the temperature difference between the two stations will have a systematic

component when the day has strong north wind and bright sunshine. In this paper, we try to use a linear model to investigate whether the hours of bright sunshine and the north wind have any effect on the temperature difference. Since the readings from both stations are random and measured with errors, an alternative estimation method rather than the method of least squares should be used. In the following sections, we will give a brief review of the measurement error model and the estimation methods. Finally, we will present the analysis and give a conclusion on the finding.

2. Measurement Error Model

Regression analysis has been studied for the past several decades extensively. In the classical case, we assume that the explanatory variables are fixed and can be observed without error. However, in many practical situations the explanatory variables can not be observed directly and they can be fixed or random. Consider the following model

$$\begin{aligned}
 (2.1) \quad & Y = X\beta + \varepsilon, & E[\varepsilon_i] = 0, E[\varepsilon_i^2] = \sigma_\varepsilon^2, \\
 & y_i = Y_i + u_i, & E[u_i] = 0, E[u_i^2] = \sigma_u^2 \\
 & x_i = \underline{X}_i + \underline{\varepsilon}_i, & E[\underline{\varepsilon}_i] = 0, E[\underline{\varepsilon}_i \underline{\varepsilon}_i^T] = \text{diag}(\sigma_1^2, \dots, \sigma_p^2) \\
 & \text{and } \varepsilon, u_i, \underline{\varepsilon}_i \text{ are uncorrelated,} \\
 & X = [\underline{X}_1, \underline{X}_2, \dots, \underline{X}_n]^T, & Y = [Y_1, \dots, Y_n]^T \\
 & \underline{X}_i \text{ is a } p \times 1 \text{ vector,} & i = 1, \dots, n
 \end{aligned}$$

where Y_i and \underline{X}_i are the true values and y_i and x_i are the observed values of Y_i and \underline{X}_i , respectively. (2.1) is the classical set up for the measurement error model or the error-in-variables model. When the the explanatory variables are fixed, we have the so called functional relationship. If they are random, we have the structural relationship. From (2.1), we can rewrite the model as

$$(2.2) \quad y_i = x_i^T \beta + \varepsilon_i^*, \quad \varepsilon_i^* = \varepsilon_i + u_i - \underline{\varepsilon}_i^T \beta.$$

People usually make inferences on β via (2.2) by using the method of least squares. The least squares estimator for β in (2.2) is

$$(2.3) \quad \hat{\beta} = (x^T x)^{-1} x^T y, \quad x = [x_1, x_2, \dots, x_n]^T.$$

If we substitute (2.2) into (2.3) and take expectation, we get

$$(2.4) \quad E[\hat{\beta}] = \beta + E[(x^T x)^{-1} x^T \epsilon^*]$$

Since x and ϵ^* are correlated, the expectation of the last term in (2.4) does not vanish. Many statisticians have pointed out that the LS estimator of β in (2.2) is biased and inconsistent, and they recommended that other estimation methods should be used to estimate β (see Fuller 1987). However, Draper and Smith (1981) have mentioned some conditions which the LS estimation can still be used in (2.2), and they are

1. When σ_k^2 , $k = 1, \dots, p$, are small compared to $\sigma_{x_k}^2$, the variance of X_k , the bias in (2.4) will be small and approaches zero as $\frac{\sigma_{x_k}^2}{\sigma_k^2} \rightarrow \infty$.
2. When the explanatory variables are fixed and controlled by the planner, the x and ϵ^* are uncorrelated.
3. We use the model

$$Y = x\beta + \epsilon.$$

In this paper, the instrumental variables estimation and the orthogonal regression estimation will be discussed in the following sections to handle the estimation of β in (2.2).

3. Instrumental variables Estimation

The instrumental variables (IV) method has been used to solve some econometric problems for many years. However, this method can also be used to handle the estimation of the regression parameters when the explanatory variables are subject to measurement error. In (2.1), assume that

there is a set of variables, W , say, which is highly correlated with X but is independent of the measurement errors e , ϵ , u . The idea of IV estimation is to extract more information about X from W , then use this extra information to estimate β . Since the measurement error u of Y and ϵ can be considered as one error term without affecting the structure of (2.1), we can rewrite (2.1) as

$$\begin{aligned} y &= X\beta + v \\ (3.1) \quad x &= X + e \end{aligned}$$

As mentioned above, we can write the instrumental variable W as

$$(3.2) \quad X = W\theta + a \quad E[a_k] = 0, \quad E[a_k a_k^T] = \sigma_a^2 I, \quad k=1, \dots, p$$

where W and a are independent.

Putting (3.2) into (3.1), we have

$$(3.3a) \quad y = W\theta\beta + v^*, \quad v^* = v + a\beta$$

$$(3.3b) \quad x = W\theta + e^*, \quad e^* = e + a$$

From (3.3), we can estimate θ from (3.3b), then use $\hat{\theta}$, y , W to estimate β by LS. The resulting estimators of θ and β are

$$\begin{aligned} \hat{\theta} &= (W^T W)^{-1} W^T x \\ \hat{\beta} &= [(W\hat{\theta})^T (W\hat{\theta})]^{-1} (W\hat{\theta})^T y \\ (3.4) \quad &= [(W(W^T W)^{-1} W^T x)^T (W(W^T W)^{-1} W^T x)]^{-1} (W(W^T W)^{-1} W^T x)^T y \\ &= [x^T W (W^T W)^{-1} W^T x]^{-1} x^T W (W^T W)^{-1} W^T y \end{aligned}$$

$$\text{or} \quad \hat{\beta} = [\hat{x}^T \hat{x}]^{-1} \hat{x}^T y, \quad \hat{x} = W\hat{\theta}$$

Substituting (3.1) into (3.4) and taking expectation, we have

$$(3.5) \quad E[\hat{\beta}] = \beta + E\{[x^T W (W^T W)^{-1} W^T x]^{-1} x^T W (W^T W)^{-1} W^T v\}.$$

Because of the independence assumption on W , the last term in (3.5) vanishes. Hence, the IV estimator of β , say $\hat{\beta}_{IV}$, is unbiased. In order to use IV estimation, we must have $q \geq p$, where $W_{n \times q}$ and $X_{n \times p}$. Also, one can use the two stage LS method to estimate β . First, one regresses x on W to get \hat{x} ; then, one regresses y on \hat{x} to get the estimate of β . For more properties of IV estimation, see Bowden and Turkington (1984).

Due to recent active research in robust regression, we try to robustify the IV estimation by using robust estimation in each stage. The type of robust estimation which we will discuss is bounded influence estimation for regression. For the model,

$$y_i = x_i^T \beta + e_i \quad E[e_i] = 0, \quad E[e_i^2] = \sigma^2,$$

the estimate of β can be obtained by solving

$$\sum_{i=1}^n \eta(x_i, r_i) x_i = 0, \quad r_i = (y_i - x_i^T \beta) / \sigma,$$

where η is the function mentioned in Hampel et. al. (1986). Wiens (1991) mentioned a simple procedure to obtain the estimate of β and its variance by using iteratively reweighted LS. This procedure can be implemented in any programming language or in any statistical package which can implement weighted LS.

4. Orthogonal Regression Estimation

In instrumental variables estimation, the estimation of β relies on the instrumental variables. When such variables do not exist, one may use the

maximum likelihood (ML) method to estimate β . We now rewrite the model

(2.1) as follows

$$(4.1) \quad Y_i = X_i^T \beta + \varepsilon_i$$

$$y_i = Y_i + u_i$$

$$x_i = X_i + e_i$$

with $\varepsilon_i \sim N(0, \sigma_\varepsilon^2)$

$$u_i \sim N(0, \sigma_u^2)$$

$$e_i \sim N(0, \Sigma_{ee}), \quad \Sigma_{ee} = \text{diag}(\sigma_1^2, \dots, \sigma_p^2)$$

$$X_i \sim N(\mu_x, \Sigma_{xx})$$

and $\varepsilon_i, u_i, e_i, X_i$ are independent.

Assume that the observation $Z_i = (y_i, x_i^T)^T$ is from a normal population with mean μ_Z and covariance Σ_{ZZ} where $\mu_Z = (\mu_x^T \beta, \mu_x^T)^T$, and

$$\Sigma_{ZZ} = \begin{bmatrix} \beta^T \Sigma_{xx} \beta + \sigma_\varepsilon^2 + \sigma_u^2 & \beta^T \Sigma_{xx} \\ \Sigma_{xx} \beta & \Sigma_{xx} + \Sigma_{ee} \end{bmatrix}.$$

Hence, the logarithm of the likelihood function is

$$(4.2) \quad \ln L = -\frac{np}{2} \ln(2\pi) - \frac{n}{2} \ln |\Sigma_{xx}| - \frac{1}{2} \text{tr} \left[\Sigma_{ZZ} \sum_{i=1}^n (Z_i - \mu_Z) (Z_i - \mu_Z)^T \right].$$

When σ_u^2 and $\sigma_k^2, k=1, \dots, p$ are known, or $\lambda_k = \frac{\sigma_u^2}{\sigma_k^2} = 1$, Kendall and Stuart

(1973), Anderson(1984), Fuller(1987) show that maximizing (4.2) is equivalent to minimizing the sum of squared perpendicular distances from the observation points to the plane defined in (4.1). In other words, the ML

estimator of β is the result from the orthogonal regression (OR) of y on x . In the least squares sense, the resulting estimator comes from minimizing the following

$$D(\beta) = \frac{\sum_{i=1}^n (y_i - x_i^T \beta)^2}{\sum_{k=1}^p \beta_k^2 + 1}.$$

Since orthogonal regression is not used in practice very often, a program for it is not included in statistical packages. However, Ammann and Van Ness (1988) have developed a routine to convert a standard regression program into an orthogonal regression program. "The idea is to use an iterative procedure to rotate the data until the regression line is horizontal. The final regression line and the data are then rotated back to the original coordinates giving the OR line".¹ Hence, the OR estimate of β for (4.1) can be obtained by using the following algorithm

1. Obtain x^0 and y^0 by subtracting \bar{x} and \bar{y} from x and y .
2. Get $\hat{\beta}^0$, by doing a regression of y^0 on x^0 . Check whether $\hat{\beta}^0$ is close to zero; if not, go to step 3.
3. Rotate the data until the plane defined by $\hat{\beta}^k$ is perpendicular to the y -axis. The rotation matrix is the matrix Q from the QR decomposition of

$$\begin{bmatrix} I_p & 0 \\ \hat{\beta}^k & 1 \end{bmatrix}.$$

¹ L. Ammann and J. Van Ness (1988)

The rotated data x^{k+1} and y^{k+1} are the columns of $[x^k | y^k] * Q$.

4. Get $\hat{\beta}^{k+1}$, by doing the regression of y^{k+1} on x^{k+1} . Check whether $\hat{\beta}^{k+1}$ is close to zero. If not, go back to step 3; otherwise, go to step 5.
5. Divide the last column, say q , of $Q^k Q^{k-1} \dots Q^0$, where Q^k is the matrix Q obtained at the k th iteration and Q^0 is the identity matrix, by the negative value of the last element of q . Form b equal to the first p elements of q^* , the resulting q from above. Also, define a equal to $\bar{y} - \bar{x}b$.
6. The OR estimate of β is giving by $[a | b^T]^T$.

The classical OR can be obtained by using the LS regression in the above procedure. If one uses the robust regression procedure instead, the robust OR will be obtained.

5. Analysis

In section 1, we have tested that the temperature at the Ellerslie station is on average higher than the temperature at the airport station. The model that we will use is

$$(5.1) \quad Y_i = \beta_0 + \beta_1 X_{1i} + \beta_2 X_{2i} + \beta_3 X_{3i}, \quad i = 1, \dots, n$$

where Y_i = True temperature at Ellerslie - true temperature at the airport;

X_{1i} = True total hours of bright sunshine at Ellerslie;

$$X_{2i} = \frac{1}{\text{True total km of north wind speed at Ellerslie}};$$

$$X_{3i} = X_{1i} * X_{2i},$$

and we will use only the days with north wind greater than 10 km per day to estimate the parameters in the model. The reason for using the temperature difference instead of the actual mean temperature is to filter out the seasonal and air-mass effects on the temperature; hence, it will leave only very local influences upon the variables of interest. Also, we expect that the temperature differences are directly proportional to the sunshine hours, but are inversely proportional to the north wind speeds. For exploration purpose, we will use the first half of the data, from 1969 - 1971; for model validation, we will use the second half of the data set, from 1972-1976. Since all readings are subject to measurement errors, the above mentioned methods are very suitable for this analysis. For the IV estimation, we will use the readings for sun and winds from Edmonton International Airport station as instrumental variables. For the OR estimation, we estimate the variance on the measurement errors by comparing the observations at Ellerslie and the airport, since on any given date the readings from both stations should be very close. Therefore, we can assume that they are fixed, but in fact they are random. Also, we assume that the measurement error for the equipment from both stations are the same and the errors are uncorrelated. If we take the differences of the same kind of reading from both stations and divide them by square root two, the measurement error variance is then calculated by the usual formula with $n-1$ replaced by n . For instance, to calculate the measurement error variance on the north wind speed, we take the readings on north wind speed from both stations. Finally, we take the differences between the two readings and divide them by square root two. For the robust case, we use a robust estimate of scale to estimate the variances. The most

common estimator is the Median Absolute Deviation (MAD) estimator, it is defined by

$$s(\omega) = 1.4826 * \text{median}(|\omega_i - \text{median}(\omega_i)|).$$

The estimated measurement error variance for the temperature is computed by the same formula.

Using the estimation methods mentioned above, we estimated the β s by four different methods. These methods were:

1. IV Estimation with the readings from the airport as the instrumental variables.
2. Robust IV Estimation with the readings from the airport as the instrumental variables. The function η that we use is

$$\eta(x_i, r_i) = w(x_i) \psi(r_i / w(x_i)),$$

$$\text{where } \psi(a) = \begin{cases} a & |a| \leq c \\ c \text{sign}(a) & |a| \geq c \end{cases}, \quad c=1.0,$$

$$\text{and } w(x_i) = (1 - h_i) / h_i^{0.5}, \quad h_i = \underline{x}_i^T (x^T x)^{-1} \underline{x}_i^T$$

3. Orthogonal Regression. We use the estimated measurement error standard deviations to scale the data, so that the λ s are all equal to one. Then we use the routine mentioned in Section 3 to compute the orthogonal regression estimates.
4. Orthogonal Regression M-estimate. We estimate the error variances by the MAD estimator and scale the data, so that the λ s are all equal to one. The β s are computed by the OR routine with the M-estimation of regression. The M-estimator for regression can be obtained by solving

$$\sum_{i=1}^n \psi(r_i) x_i = 0, \quad r_i = (y_i - x_i^T \beta) / \sigma.$$

The M-estimator for regression can be obtained by using the algorithm for bounded influence regression estimate, such as one mentioned in Section 3. We use the same ψ function as in method 2.

All the computation are performed by Proc IML of SAS; the final estimates from the above methods are listed in Table 5.1. In order to see the effect on the estimates from the least squares estimation, we also include the results from the least squares estimation in the last column in Table 5.1.

From Table 5.1, both IV estimations result in similar estimates for the β s. Also, the estimates have the right sign as we expected. For the least squares estimation, the estimates are not the same as both IV estimates because of the bias mentioned in section 2. For the Orthogonal Regression estimates, the final estimates of the β s are quite different from the results from both IV estimations. This problem may arise from the outliers in the data and the estimated measurement error variances. It is well known that the orthogonal regression estimate is unstable due to unit change in variables, such as from inches to yards. Also the classical OR estimate is very sensitive to outliers; even with a small amounts of outliers, the classical OR estimate can be completely changed. The outliers in the data can be noticed from the differences between the classical IV estimates and the robust IV estimates, and from the remote points on the residuals plots. Although the robust OR estimation can down weight the effect of the outliers in the data, the estimates from this method also rely on the accurately estimated measurement error variances and hence the λ s. Lakshminarayanan and Gunst(1984) have shown that when the λ s are incorrectly specified the classical OR estimate is inconsistent and the asymptotic variance for the β s will not be valid, because

the estimator has no finite sample moments. We think that the incorrectly estimated λ s are from the wind speeds and this is the major reason for the resulting estimates. Although we have replications on the wind speeds, the readings are taken out from different equipment. In Ellerslie station, the equipment which take the wind speeds only reads the speed from 8 directions, i.e. N, NE, E, SE, S, SW, W, NW. For the airport station, before 1971, the equipment read the speed from 16 directions; after 1971, the equipment can read the speed from 36 directions. This difference could cause the unequal variance from the equipment, and hence violate our assumption of equal error variances from both stations. Also, the total wind speeds on each day are computed by summing up the 24 hourly readings. Therefore, as the wind direction can change within an hour, the equipment from the airport station will produce more accurate reading than the equipment in the Ellerslie station. This may also cause the incorrectly estimated of the λ s. The misspecification of the measurement error structure may also cause a problem in the OR estimations. At the beginning, we assume that the measurement errors are uncorrelated. When the interaction terms are added, the covariance matrix of the measurement error will not be diagonal. When this is true, we have to employ the generalized version of the OR and the simple routine mentioned can not be used. In the robust case, the estimates can be obtained by directly solving the estimating equations but this could lead to complicated computation.

Due to the above mentioned problems in the OR estimations and the least squares estimation, the rest of the analysis will be based on both IV estimations. Since β_0 is the most insignificant from both IV estimations, we will delete it from (5.1) and estimate $\beta_1, \beta_2, \beta_3$ again from the following model

$$(5.2) \quad Y_i = \beta_1 X_{1i} + \beta_2 X_{2i} + \beta_3 X_{3i}, \quad i = 1, \dots, n.$$

Both IV estimations are listed in Table (5.2). Again, both IV estimations result in similar estimates for all parameters except for β_2 ; we think that the outliers in the data may have the effect on it. Also, both IV estimations have shown that β_1 , and β_3 are insignificant; hence, we will remove them from (5.2) in turn. First, we remove X_1 from (5.2), since it is the most insignificant in the model. After X_1 is removed, β_2 still remains significant in the new model and β_3 also remains insignificant. Finally, we remove both X_1 and X_3 from (5.2) and get the following model

$$(5.3) \quad Y_i = \beta_2 X_{2i}, \quad i = 1, \dots, n.$$

The final results are presented in Table 5.3. Once again, the final estimates from both IV estimations are quite similar, and the residuals plots do not have any unusual trends. For the model validation, we estimate β_2 in (5.3) by using the second half of the data, from 1972 - 1976. The final estimates and the 95% confidence intervals are presented in Table 5.3. Looking at the estimates for both data sets, we can see that the estimates of β_2 in the second half of the data set is about 2 units higher than the estimates of β_2 in the first half from robust IV estimation. This might cause by the height of the trees, since the trees do increase in height as age increases. As a result, the trees may cause the effect of the north wind, i.e. β_2 , to be higher in the second half of the data. However, the 95% confidence intervals produced from both data sets do overlap more than 50%. This indication suggests that β_2 in (5.3) is similarly estimated by the IV estimations. Hence, we conclude that (5.3) is the final model to study the effect of the temperature differences.

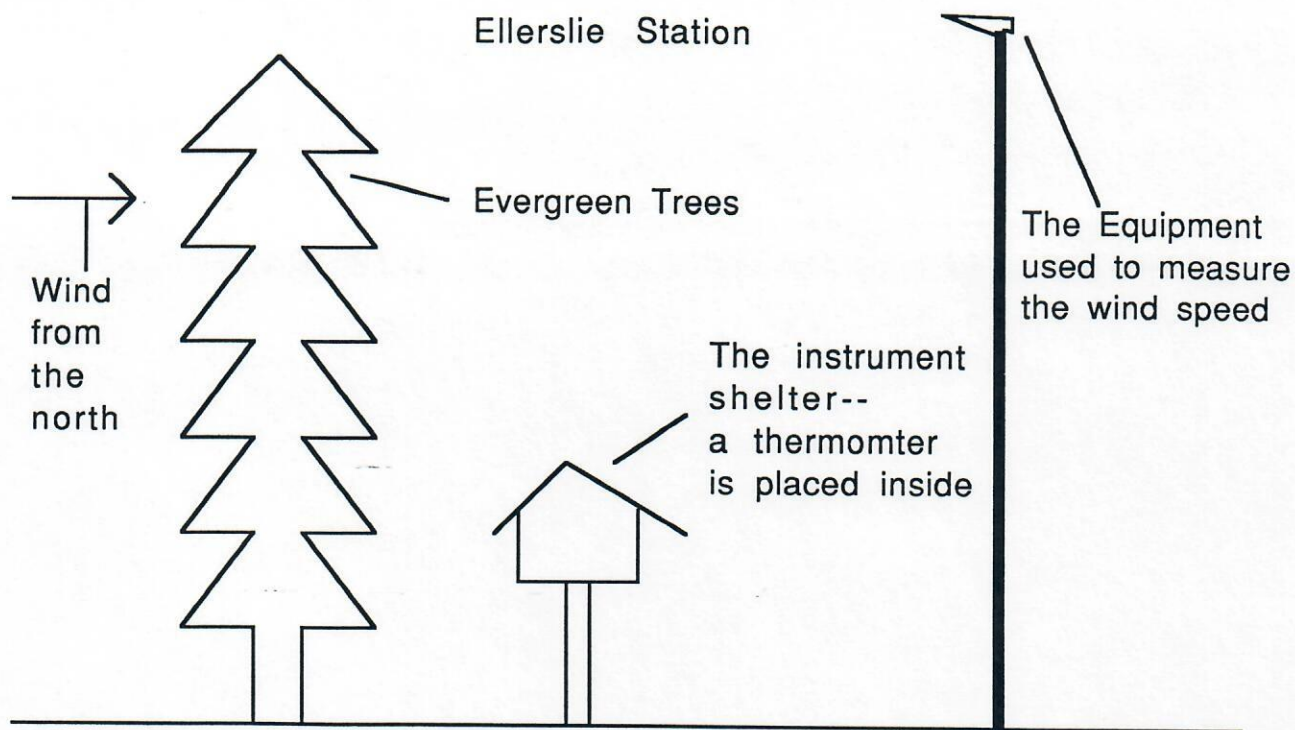
6. Conclusion

In the previous section, we have concluded that (5.3) is our final model to study the temperature differences. Based on this model, we can conclude

that the north winds do have significant effect on the temperature differences, and hence the shelterbelt does cause higher temperature at the Ellerslie station. Also we can see that the temperature differences is more significant when the north winds is moderate, and the differences reduce as the north winds getting stronger. Finally, we want to make a remark about the data set that we used. Since the shelter effect will work more efficiently when the place has bright sunshine and moderate to strong north wind altogether, hourly readings of the variables show when this condition is obtained. Hence, if one can obtain the hourly readings on the temperature, the north wind, and the bright sunshine, a more detailed analysis can be performed.

REFERENCES

- Ammann, L. and Van Ness, J. (1988). A Routine for Converting Regression Algorithms into Corresponding Orthogonal Regression Algorithms. *ACM Transaction Mathematical Software*, 14, 76-87.
- Anderson, T. W. (1984). Estimating Linear Statistical Relationships. *Annals of Statistics*, 12, 1-45.
- Bowden, R. J. and Turkington, D. A. (1984). *Instrumental Variables*, 1st ed., New York, Cambridge University Press.
- Draper, N. R. and Smith, H. (1981). *Applied Regression Analysis*. 2nd ed., New York, John Wiley.
- Fuller, W. A. (1987). *Measurement Error Models*, 1st ed., New York, John Wiley.
- Hampel, F. R., Ronchetti, R., Rousseeuw, P. J. and Stahel, W. (1986). *Robust Statistics: The Approach Based on The Influence Function*, New York, John Wiley.
- Kendall, M. G. and Stuart, A. (1971), *The Advanced Theory of Statistics*, Vol. 2, 4th ed., New York, Hafner.
- Lakshminarayanan, M. Y. and Gunst, R. F. (1984). Estimation of Parameters in Linear Structural Relationships: Sensitivity to the Choice of the Ratio of Error Variances, *Biometrika*, 71, 569-573.
- Wiens, D. P. (1991). A Note on the Computation of Robust, Bounded Influence Estimates and Test Statistics in Regression, *Computational Statistics and Data Analysis*, to appear.



The Edmonton International Airport Station

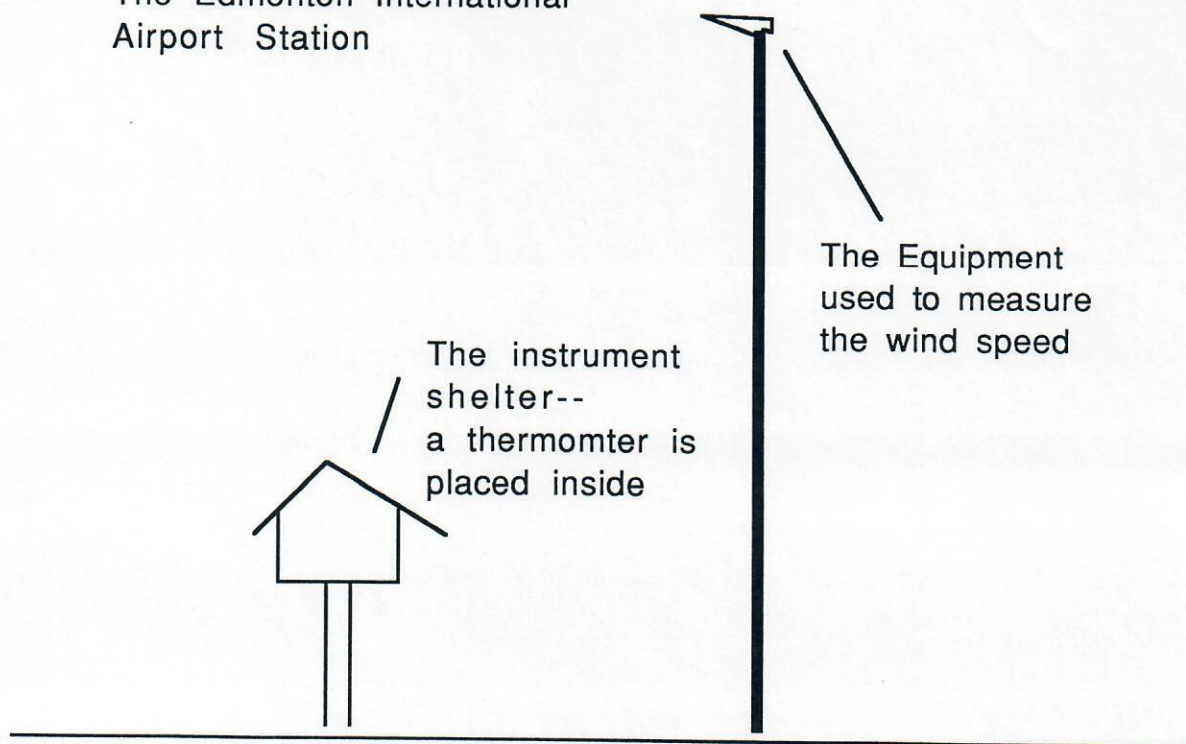


Figure 1.1

Table 5.1
The estimates of the β s in model (5.1)

	Method 1	Method 2	Method 3	Method 4	Least squares
β_0	-0.85989 (0.58760)	-0.85535 (0.58462)	87.6228 (580.847)	-660.489* (0.04647)	0.41145* (0.15705)
β_1	0.16251 (0.08946)	0.14594* (0.05870)	-11.3256 (75.2676)	64.9176* (0.01296)	0.00123 (0.01971)
β_2	48.8421* (22.0203)	48.9249* (20.6543)	-3318.92 (22101.6)	28831.0* (0.00112)	-0.16581 (4.32306)
β_3	-5.75715 (3.38392)	-5.22093* (2.09064)	430.833 (2859.59)	-3032.62* (0.00914)	0.46750 (0.55420)

NOTE: * At 5% level of significance, the hypothesis $H_0 : \beta_j = 0$ is rejected.
The value inside the parenthesis is the estimated standard deviation.

FIGURE 5.1
RESIDUALS PLOT FOR CLASSICAL IV ESTIMATION
FOR MODEL (5.1)

Plot of RESIDUAL*PREDICT. Legend: A = 1 obs, B = 2 obs, etc.

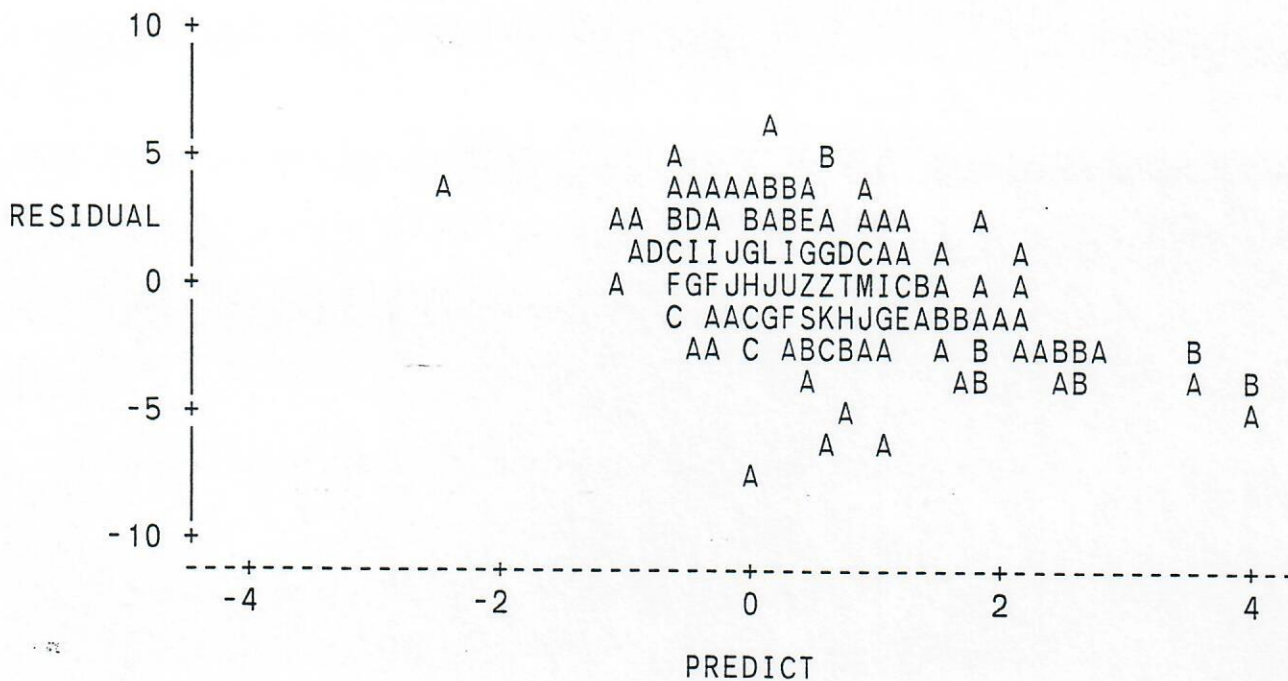


FIGURE 5.2
RESIDUALS PLOT FOR CLASSICAL IV ESTIMATION
FOR MODEL (5.1)

Plot of RESIDUAL*X1. Legend: A = 1 obs, B = 2 obs, etc.

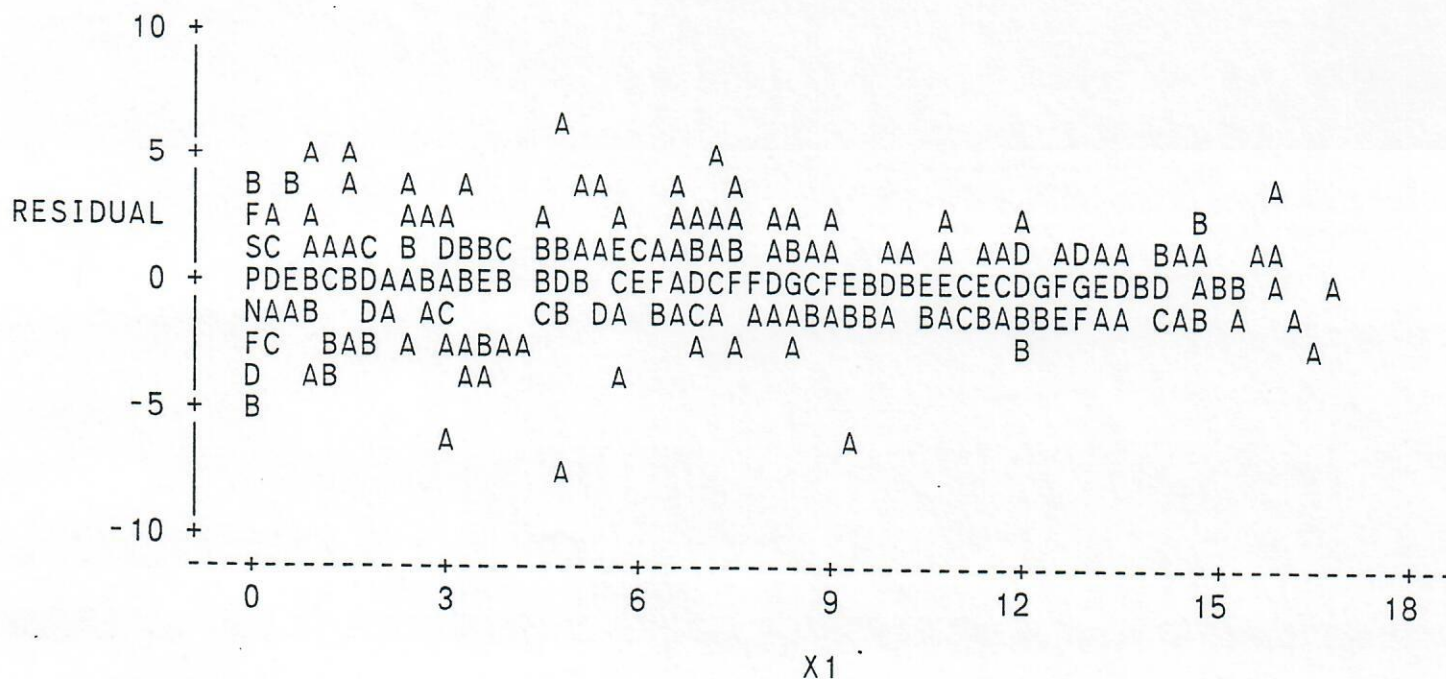


FIGURE 5.3
RESIDUALS PLOT FOR CLASSICAL IV ESTIMATION
FOR MODEL (5.1)

Plot of RESIDUAL*X2. Legend: A = 1 obs, B = 2 obs, etc.

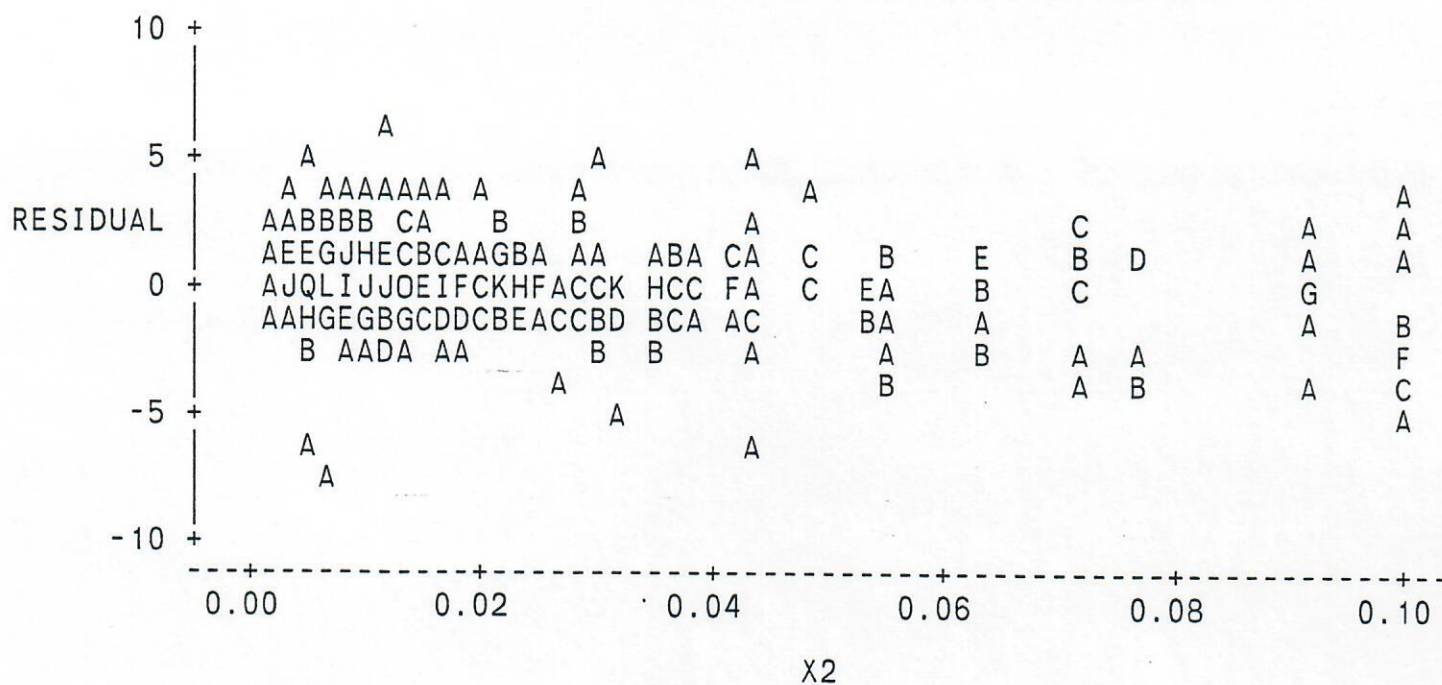


FIGURE 5.4
RESIDUALS PLOT FOR CLASSICAL IV ESTIMATION
FOR MODEL (5.1)

Plot of RESIDUAL*X3. Legend: A = 1 obs, B = 2 obs, etc.

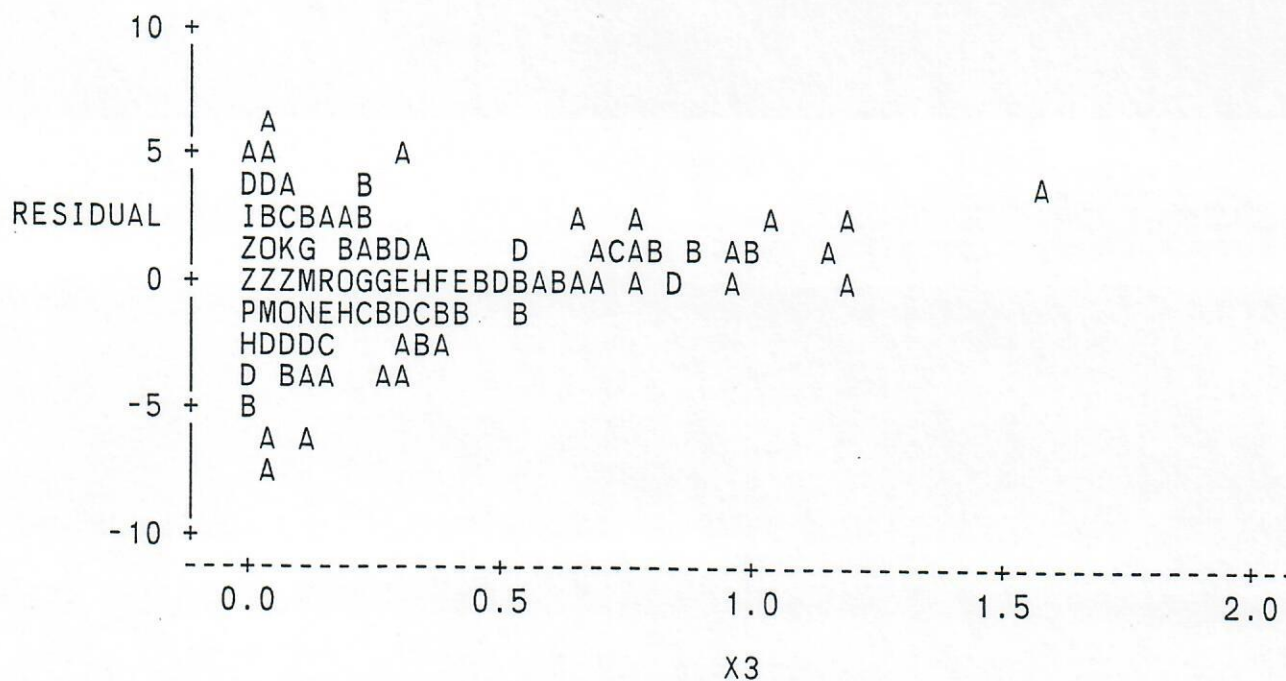


FIGURE 5.5
RESIDUALS PLOT FOR ROBUST IV ESTIMATION
FOR MODEL (5.1)

Plot of RESIDUAL*PREDICT. Legend: A = 1 obs, B = 2 obs, etc.

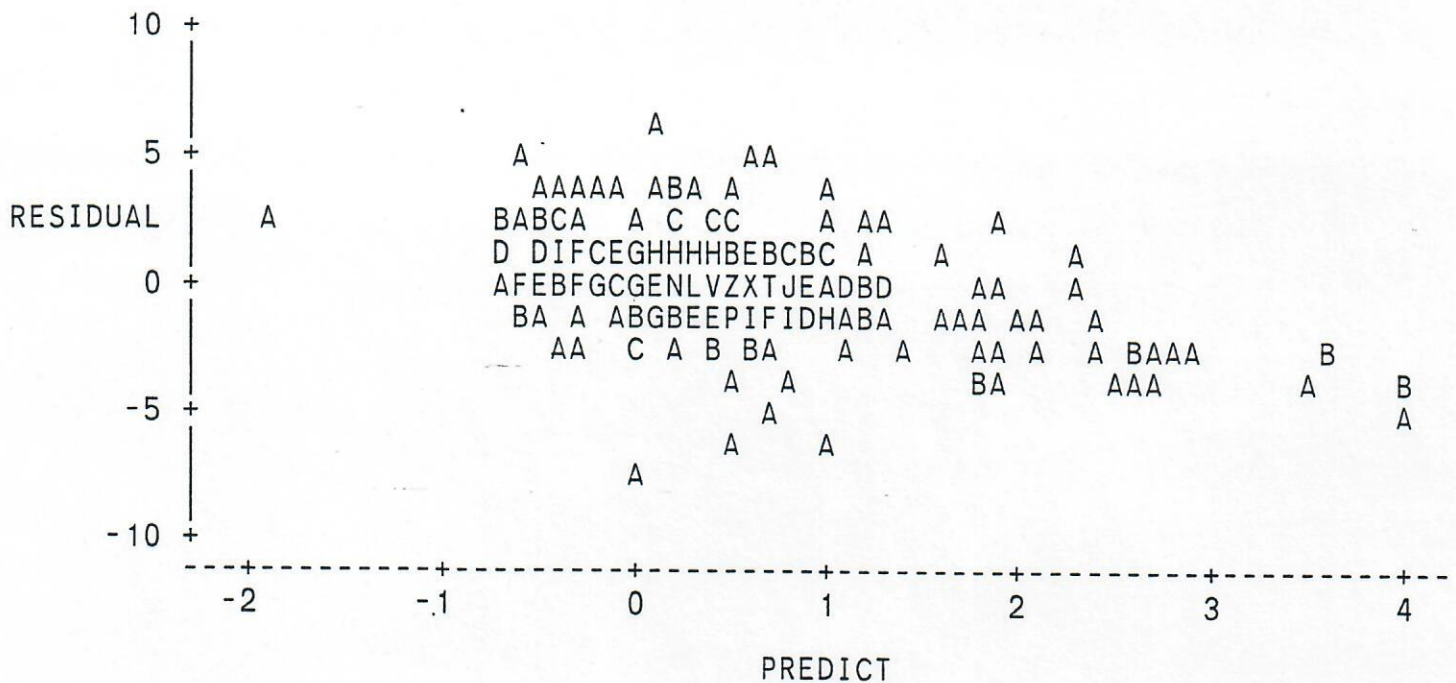


FIGURE 5.6
RESIDUALS PLOT FOR ROBUST IV ESTIMATION
FOR MODEL (5.1)

Plot of RESIDUAL*X1. Legend: A = 1 obs, B = 2 obs, etc.

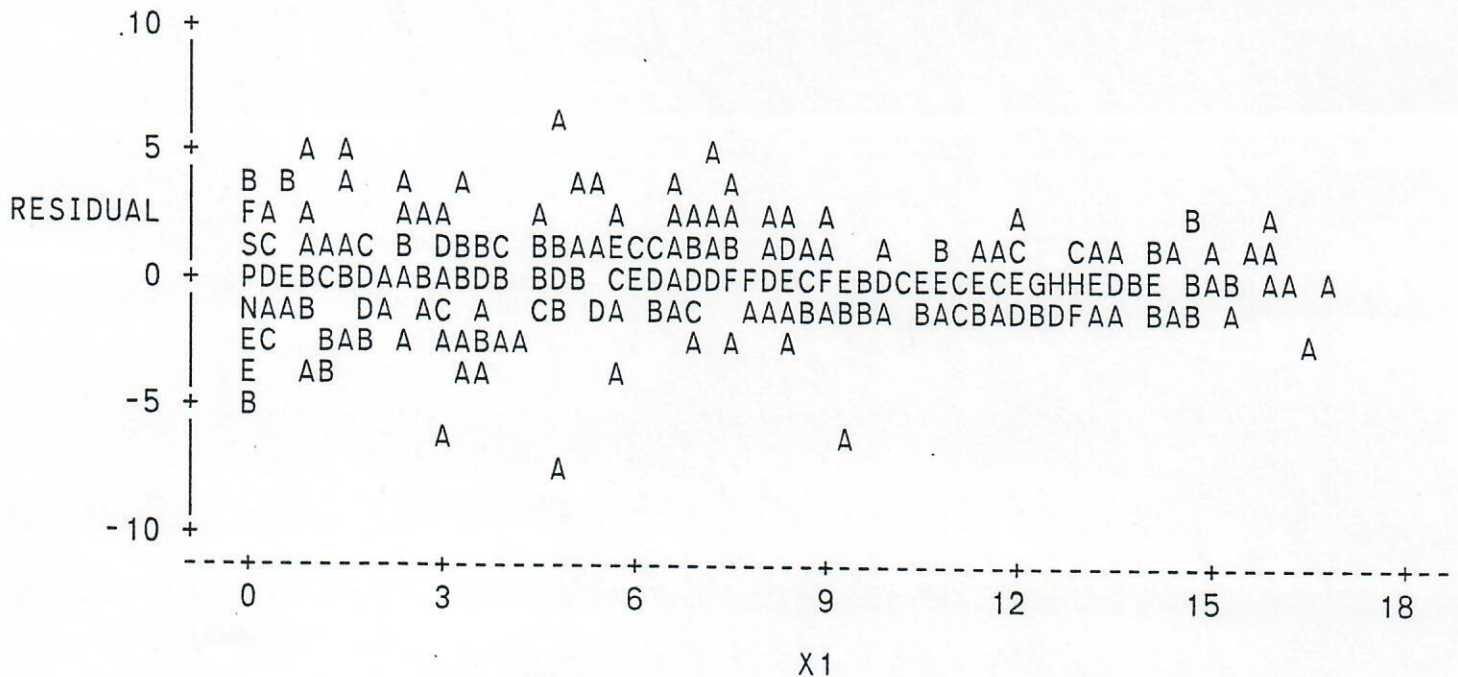


FIGURE 5.7
RESIDUALS PLOT FOR ROBUST IV ESTIMATION
FOR MODEL (5.1)

Plot of RESIDUAL*X2. Legend: A = 1 obs, B = 2 obs, etc.

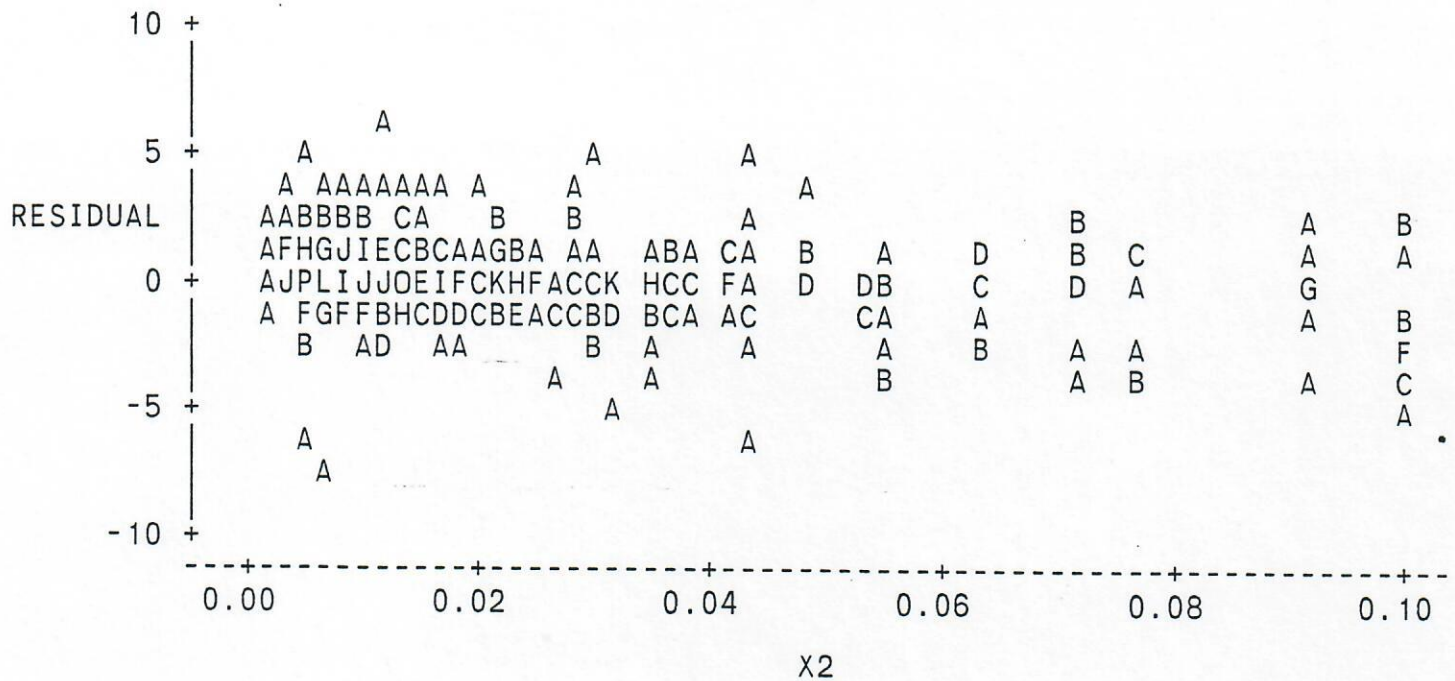


FIGURE 5.8
RESIDUALS PLOT FOR ROBUST IV ESTIMATION
FOR MODEL (5.1)

Plot of RESIDUAL*X3. Legend: A = 1 obs, B = 2 obs, etc.

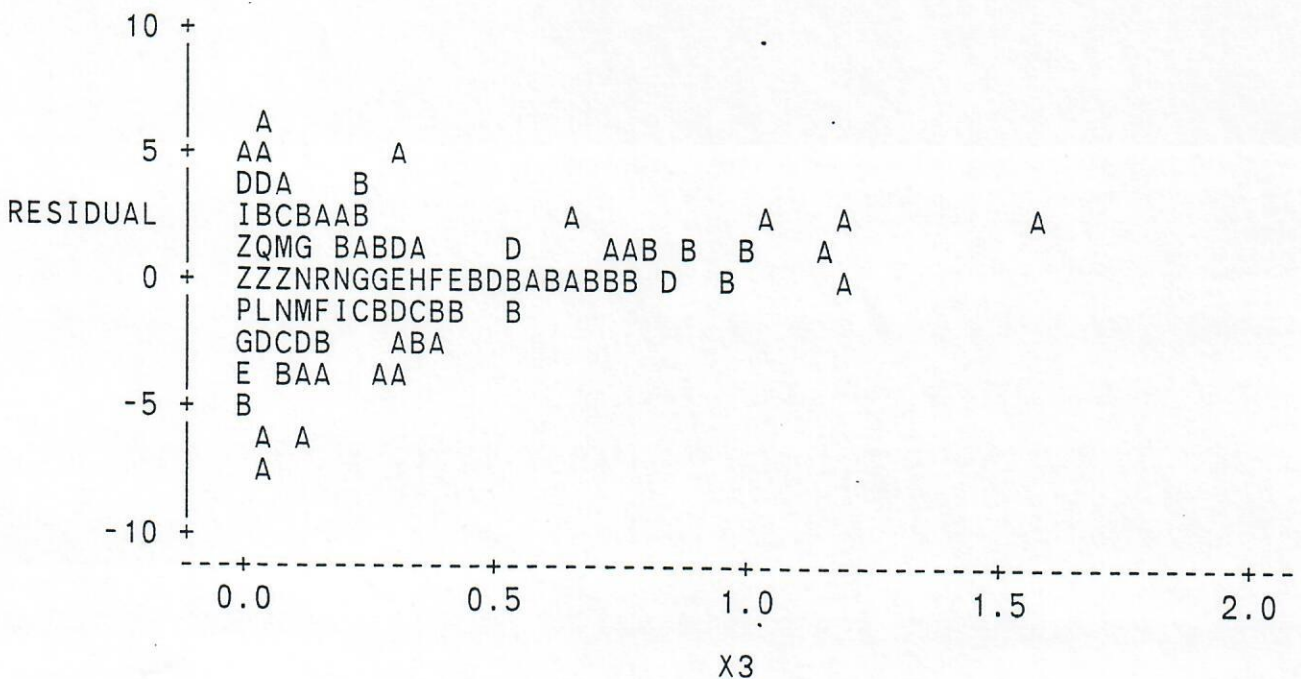


Table 5.2
The estimates of the β s in model (5.2)

	Method 1	Method 2
β_1	0.06541 (0.05353)	0.06733 (0.03836)
β_2	17.3768* (4.23944)	21.8559* (4.77357)
β_3	-2.17549 (2.08502)	-2.72308 (1.41718)

NOTE: * At 5% level of significance, the hypothesis $H_0 : \beta_j = 0$ is rejected.
The value inside the parenthesis is the estimated standard deviation.

TABLE 5.3
The estimates of the β s in model (5.3)

		YEAR 1967-1971		YEAR 1972-1976	
		Method 1	Method 2	Method 1	Method 2
β_2	LOWBOUND	13.7744	13.5950	17.0282	15.7730
	ESTIMATE	18.7341	18.5594	22.4745	20.9683
	UPBOUND	23.6938	23.5238	27.9207	26.1637
σ		1.44474	1.01939	1.45171	0.96081

NOTE : LOWBOUND --- 95% CONFIDENCE INTERVAL LOWER BOUND.
UPBOUND --- 95% CONFIDENCE INTERVAL UPPER BOUND.

FIGURE 5.9
RESIDUALS PLOT FOR CLASSICAL IV ESTIMATION
FOR MODEL (5.3), YEAR 1967-1971

Plot of RESIDUAL*PREDICT. Legend: A = 1 obs, B = 2 obs, etc.

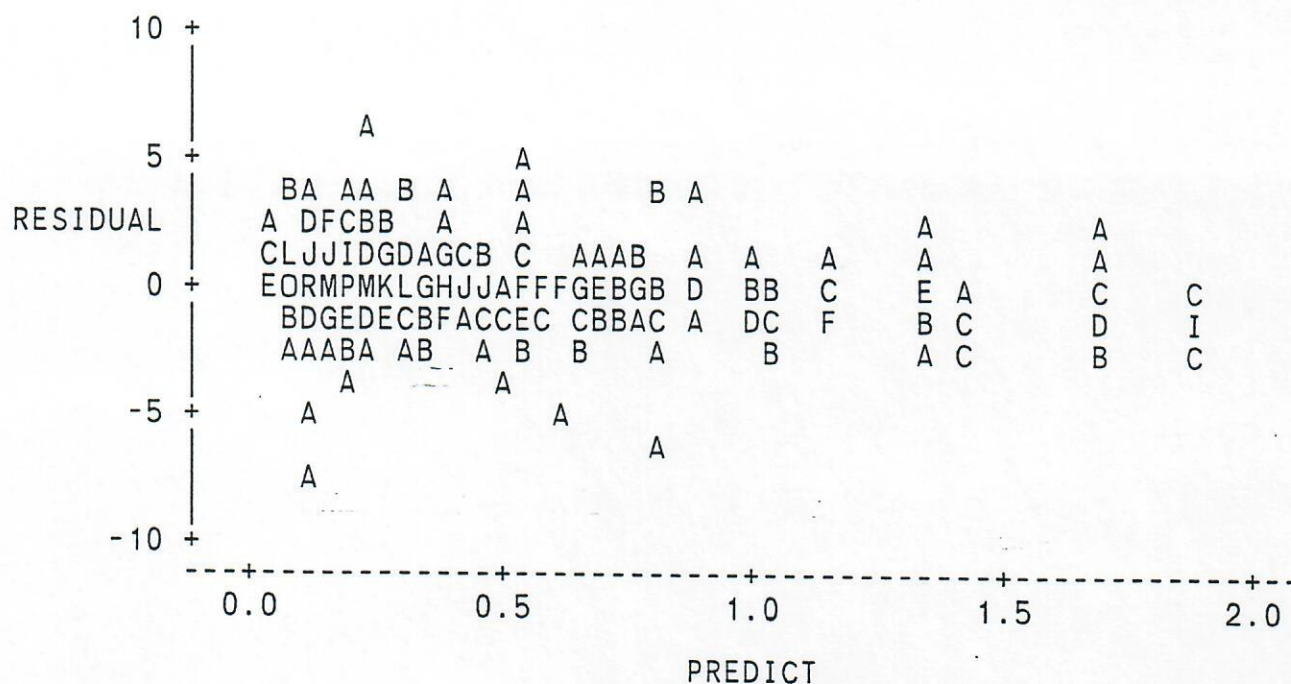


FIGURE 5.10
RESIDUALS PLOT FOR CLASSICAL IV ESTIMATION
FOR MODEL (5.3), YEAR 1967-1971

Plot of RESIDUAL*X2. Legend: A = 1 obs, B = 2 obs, etc.

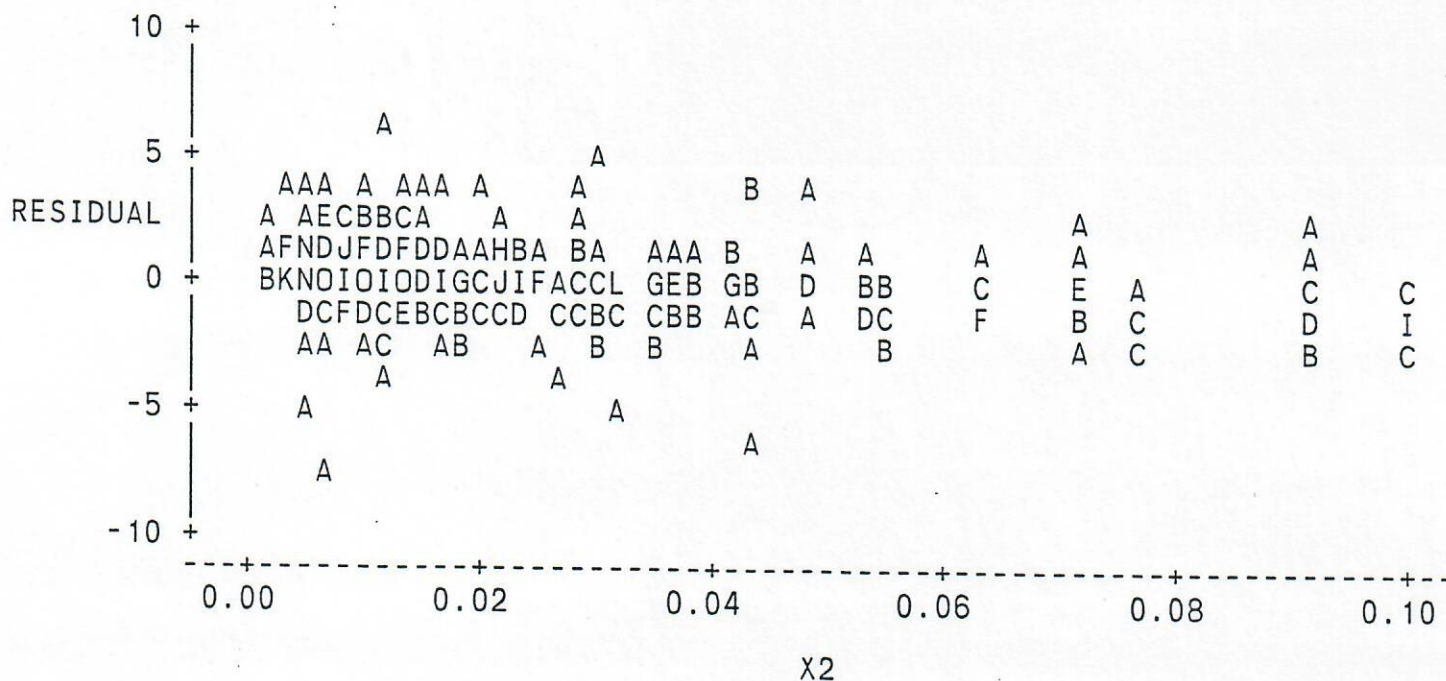


FIGURE 5.11
RESIDUALS PLOT FOR ROBUST IV ESTIMATION
FOR MODEL (5.3), YEAR 1967-1971

Plot of RESIDUAL*PREDICT. Legend: A = 1 obs, B = 2 obs, etc.

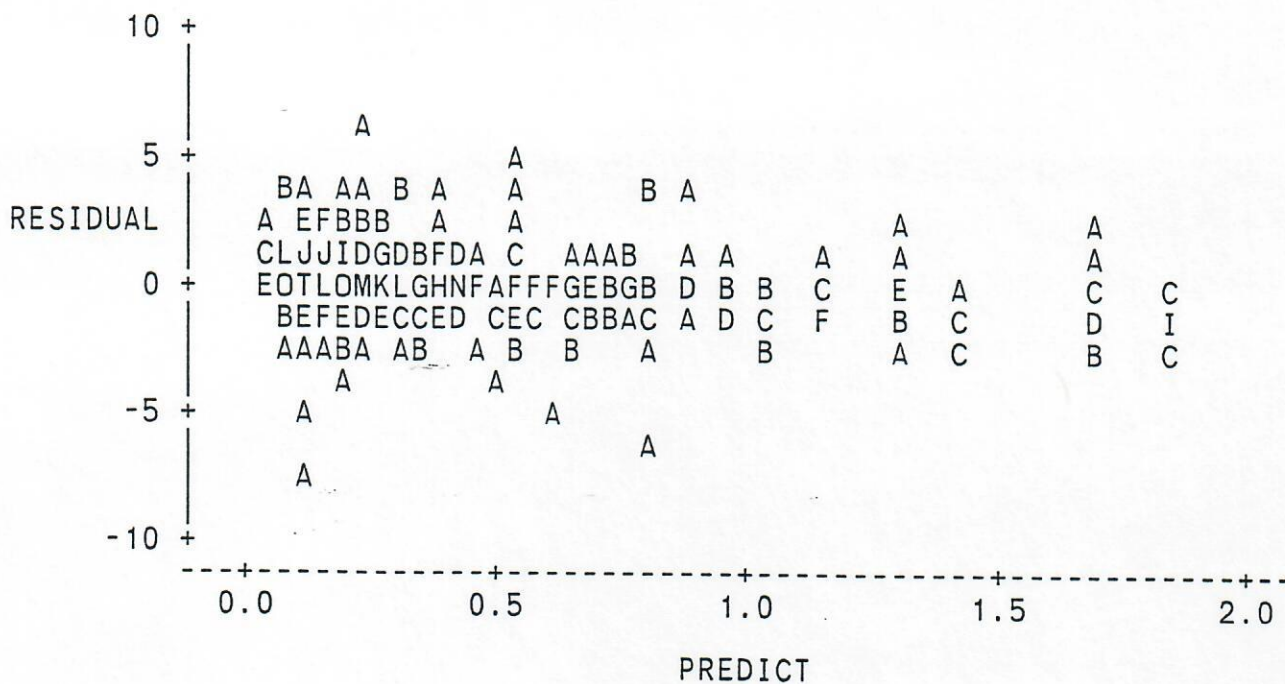


FIGURE 5.12
RESIDUALS PLOT FOR ROBUST IV ESTIMATION
FOR MODEL (5.3), YEAR 1967-1971

Plot of RESIDUAL*X2. Legend: A = 1 obs, B = 2 obs, etc.

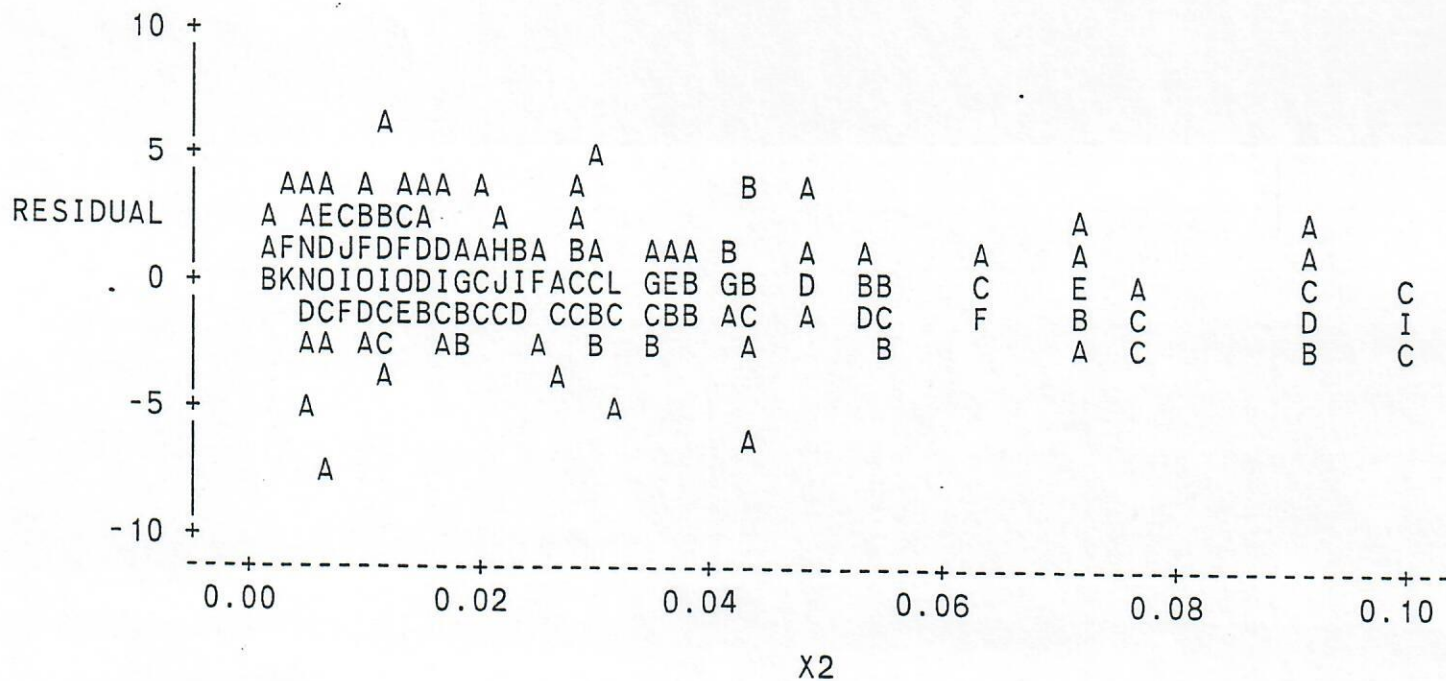


FIGURE 5.13
RESIDUALS PLOT FOR CLASSICAL IV ESTIMATION
FOR MODEL (5.3), YEAR 1972-1976

Plot of RESIDUAL*PREDICT. Legend: A = 1 obs, B = 2 obs, etc.

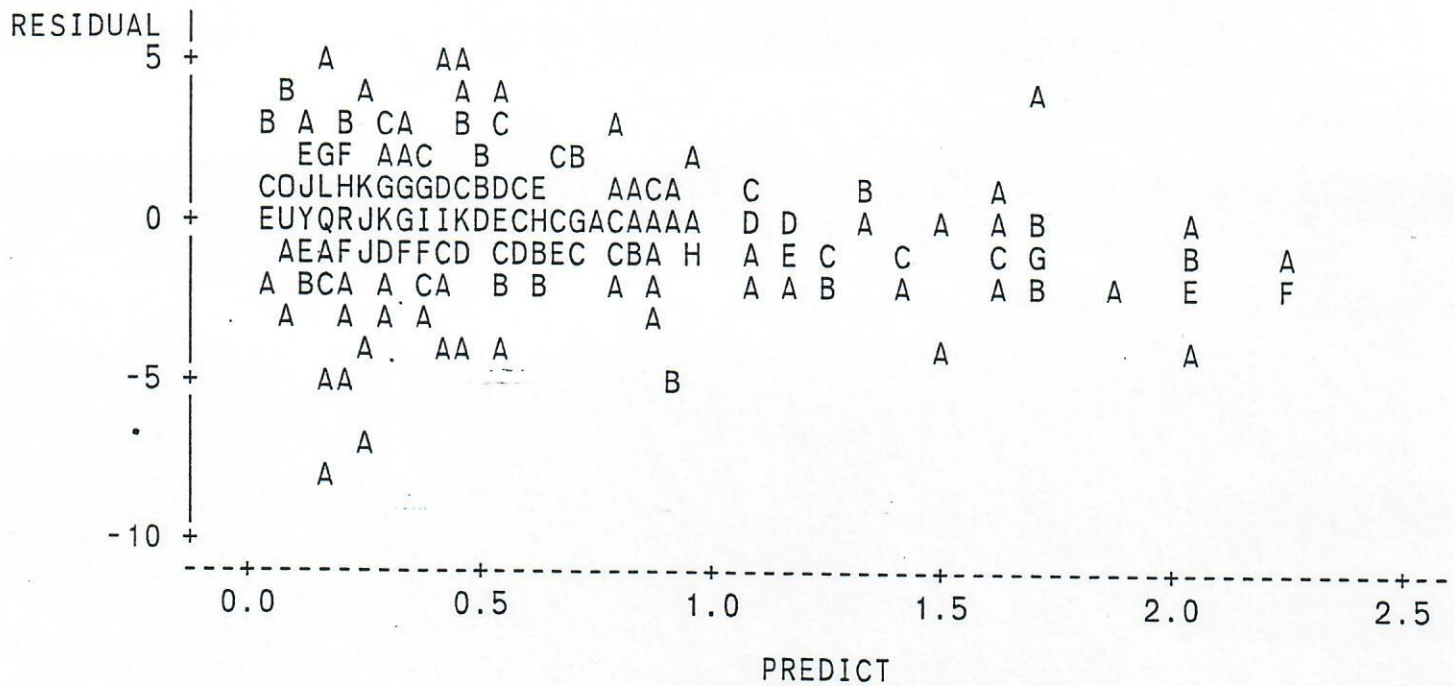


FIGURE 5.14
RESIDUALS PLOT FOR CLASSICAL IV ESTIMATION
FOR MODEL (5.3), YEAR 1972-1976

Plot of RESIDUAL*X2. Legend: A = 1 obs, B = 2 obs, etc.

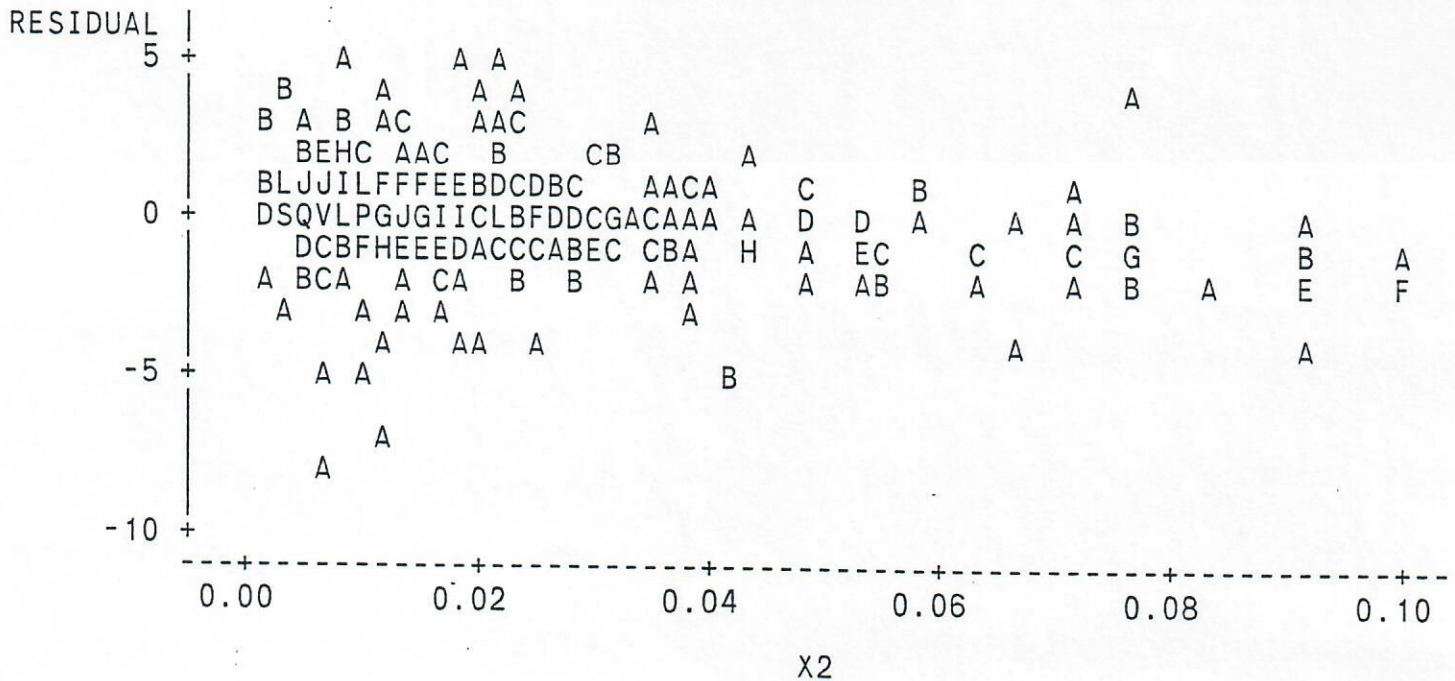


FIGURE 5.15
RESIDUALS PLOT FOR ROBUST IV ESTIMATION
FOR MODEL (5.3), YEAR 1972-1976

Plot of RESIDUAL*PREDICT. Legend: A = 1 obs, B = 2 obs, etc.

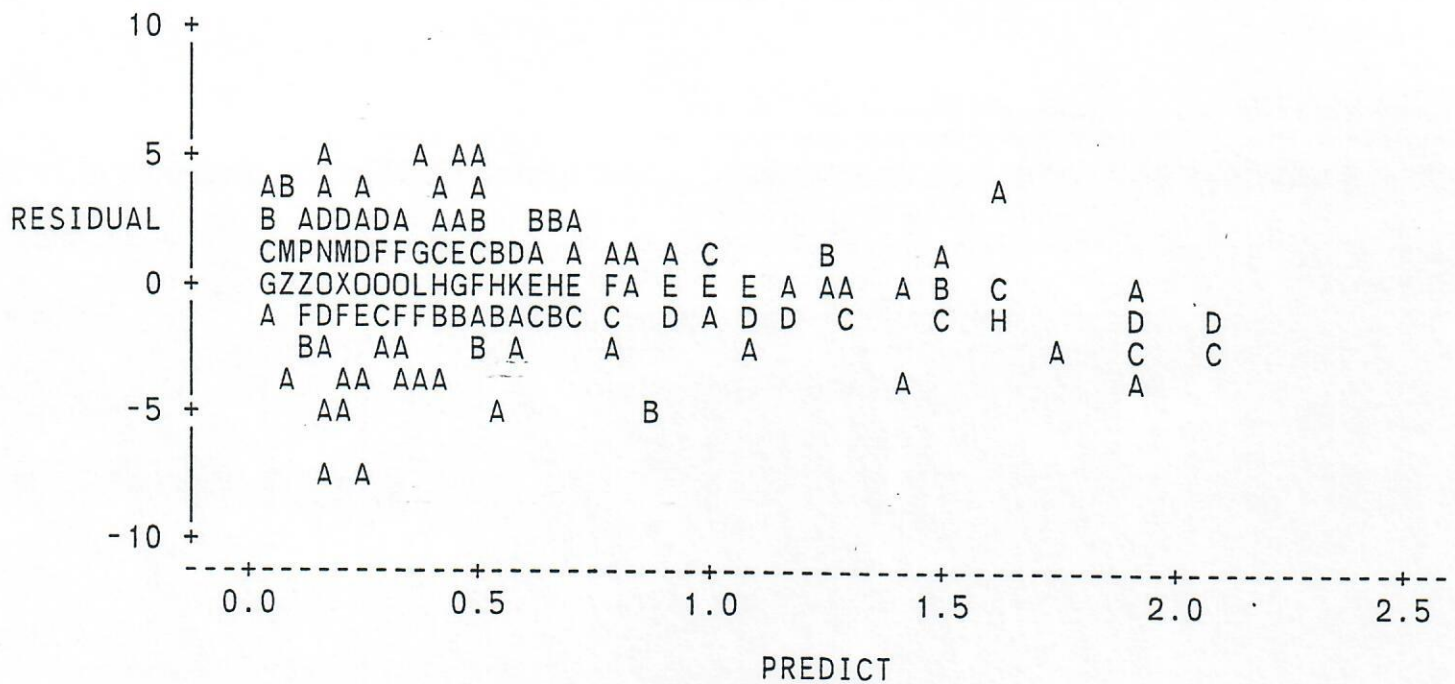


FIGURE 5.16
RESIDUALS PLOT FOR ROBUST IV ESTIMATION
FOR MODEL (5.3), YEAR 1972-1976

Plot of RESIDUAL*X2. Legend: A = 1 obs, B = 2 obs, etc.

