DOCUMENT RESUME

ED 446 142 TM 031 904

AUTHOR	Thompson, Russel L.
TITLE	An Introduction to Graphical and Mathematical Methods for Detecting Heteroscedasticity in Linear Regression.
PUB DATE	2000-01-000
NOTE	45p.; Paper presented at the Annual Meeting of the Southwest Educational Research Association (Dallas, TX, January 27-29, 2000).
PUB TYPE	Reports - Descriptive (141) -- Speeches/Meeting Papers (150)
EDRS PRICE	MF01/PC02 Plus Postage.
DESCRIPTORS	*Graphs; *Least Squares Statistics; *Regression (Statistics)
IDENTIFIERS	*Heteroscedasticity (Statistics); *Homoscedasticity (Statistics)

ABSTRACT

Homoscedasticity is an important assumption of linear regression. This paper explains what it is and why it is important to the researcher. Graphical and mathematical methods for testing the homoscedasticity assumption are demonstrated. Sources of homoscedasticity and types of homoscedasticity are discussed, and methods for correction are demonstrated. Graphs are used to illustrate different patterns that may be caused by heteroscedasticity. An extensive example for using Weighted Least Squares regression is provided using both the Statistical Package for the Social Sciences (SPSS) and a step-by-step manual process. SPSS code for reproducing all examples is included. Finally, examples are used to highlight the interactive relationship between good experimental design and sound statistical practice. Appendixes contain the listing for residual plot examples and the listing for correction examples. (Contains 4 tables, 17 figures, and 22 references.) (Author/SLD)

Running head: DETECTING HETEROSCEDASTICITY

An Introduction to Graphical and Mathematical Methods for Detecting Heteroscedasticity in Linear Regression

Russel L. Thompson

Texas A&M University 77843-4225

Paper presented at the annual meeting of the Southwest Educational Research Association, Dallas, TX, January 27-29, 2000. Special thanks to Dan Johnson for initial conversations that lead to my interest in this topic. Special thanks to Bruce Thompson for valuable comments on this paper.

Abstract

Homoscedasticity is an important assumption of linear regression. This paper explains what it is and why it is important to the researcher. Graphical and mathematical methods for testing the homoscedasticity assumption are demonstrated. Sources of homoscedasticity and types of homoscedasticity are discussed, and methods for correction are demonstrated. Graphs are used to illustrate different patterns that may be caused by heteroscedasticity. An extensive example for using Weighted Least Squares (WLS) regression is provided using both SPSS and a step-by-step, manual process. SPSS code for reproducing all examples is included. Finally, examples are used to highlight the interactive relationship between good experimental design and sound statistical practice.

An Introduction to Graphical and Mathematical

Methods for Detecting Heteroscedasticity in Linear Regression

The purpose of this paper is to discuss homoscedasticity, an important assumption of linear regression. Examples will demonstrate how good experimental design can improve the chances that analysis will be conducted correctly and will demonstrate how examining statistical assumptions can inform the interpretation of experimental results.

Many of the methods presented here invoke graphical techniques. The recent report of the APA Task Force on Statistical Inference, published in the August, 1999 issue of the American Psychologist, emphasized the value and importance of using graphics to understand and communicate data dynamics, and especially to evaluate methodological assumptions such as homoscedasticity.

All of the examples in this paper involve a single predictor variable and dependent variable in order to make the discussion of these concepts more accessible. However, the reader should keep in mind that the concepts discussed in this paper can be generalized to analyses with multiple predictor variables.

Review of Linear Regression

Linear regression is a process of creating a best-fit line between a set of predictor variables (PVs) and a single dependent variable (DV). If the assumptions are met, Ordinary Least Squares (OLS) regression results in minimized squared distances between the measured points of the PV and the predicted points of the regression line. An equation that describes the best-fit line between k predictor variables (X_1 through X_k) and a dependent variable (Y) can be written as $Y = \beta_1(Z_{X1}) + \beta_2(Z_{X2}) + \ldots \beta_k(Z_{Xk}) + e$. The error term (e), also called the residual, represents the difference between the value predicted by the equation (Yhat) and the value of the dependent variable (Z_Y). Yhat and e are synthetic variables as opposed to measured variables $X_{1..j}$ and Y. Yhat is the focus of the analysis as explained by Thompson (1992).

The beta weights ($\beta_1, \beta_2, \ldots \beta_k$) are applied to the standardized form of their respective PVs to estimate the DV. The beta weight for a given PV will be zero if that PV contributes nothing either directly to the prediction of the DV, in the context of the presence of the full set of predictors. As Thompson (1998) noted,

> The weights can be greatly influenced by which variables are included or are excluded from a given analysis... Any interpretations of weights must be considered *context-specific*. Any change in the variables in the model can radically alter all of the weights. Too few researchers appreciate the potential magnitudes of these impacts. (p. 23, emphasis in original)

Obviously, this means that beta weights to some degree reflect in part the predictive value of the PVs.

However, this truth unfortunately has led to two common misinterpretations of beta weights: (a) the misconception that a beta weight directly and exclusively measures the "relationship" beteen a PV and the DV, and (b) the misconception that a PV with a beta weight of zero has no predictive value. Instead, a beta weight indicates the number of units that the DV is predicted to change for exactly one unit of change in a given PV; this statement is entirely specific to the context of a given DV and a given set of predictors, and will not generalize to adding or deleting even only a single PV.

Structure coefficients are also important in regression research (Thompson & Borrello, 1985). Structure coefficients are correlation coefficients between the predicted DV scores (i.e., the Yhat scores) and each PV. Because the multiple $\underline{R}$ actually equals the $\underline{r}$ between $\underline{Y}_i$ and Yhat_i, Yhat is actually the focus of the regression analysis (Thompson, 1998, 1999). Structure coefficients are important in regression, because they reveal the alternative "structures" underlying one of the variables actually being evaluated in the analysis. It is only by interpreting both beta weights and structure coefficients (or beta weights and zero-order correlations between the PVs and the DV) that one can fully understand the direct and indirect contributions of the PVs.

The degree to which the regression line (or plan or hyperplane) fits the actual data is termed goodness-of-fit and is traditionally expressed as R^2. Multiple R^2 can be used as an effect size that quantifies the degree to which the regression equation fits the actual data. It is also possible, and usually desirable, to compute an adjusted R^2 that incorporates information about the sample size, number of variables, and population or future sample effect size to give a more accurate reflection of the magnitude of the relationship (Snyder & Lawson, 1993).

Assumptions of Linear Regression

Berry and Feldman (1985) list seven assumptions for linear regression that "must be met to be able to appropriately estimate the population parameters and conduct tests of statistical significance" (p. 10). The following assumptions are required for linear regression: (a) all variables are measured without error and at the interval level; (b) the mean value of the residuals is 0; (c) the variance of the residuals is constant, (i.e., homoscedasticity); (d) the residuals are uncorrelated with one another; (e) the residuals are uncorrelated with the predictor variables; (f) none of the predictor variables is perfectly correlated linearly with any of the other predictor variables; and (g) the residual terms are normally distributed. This paper addresses the third assumption, homoscedasticity.

The homoscedasticity assumption is more formally stated as $VAR(e_j) = \sigma^2$, that is, the variance of the error or residual term at each point j is equal to the variance for all residuals. The

Gauss-Markov theorem states that when the seven methodological assumptions listed above are met, the least squares estimators of regression parameters are unbiased and efficient. In short hand, the least squares estimators are said to be BLUE: Best Linear Unbiased Estimators.

Homoscedasticity in Linear Regression

Homoscedasticity describes the consistency of variance of the error term e (residual) at each level of the predictor variable. To simplify the discussion, assume a regression equation for a single predictor variable: $Z_Y = \beta(Z_X) + e$. When the assumption of homoscedasticity is met, the spread-out-ness (variance) of the error term e is the same for all values of X. The term used to describe data that violate this condition is *heteroscedasticity.* When heteroscedasticity occurs in the data for one or more points of the predictor variable, the linear regression line no longer provides a uniformly best fit for all data points throughout the distribution.

Guilford and Fruchter (1978) explain homoscedasticity in terms of the standard error of estimate (of the regression line). The standard error of estimate is an index of the variance of measured values around each predicted value. The smaller the variance, the more accurate the prediction. The standard error of estimate can be used to quantify the margin of error of prediction, but only if the range of observed values are "fairly uniform" (p. 301). Imagine a series of normal curves, one for each predicted value, oriented horizontally with their means centered on the regression line with each occurrence of a measured value at the predicted value plotted on the curve in a histogram. The histogram of measured values should lie within the normal curve to satisfy homoscedasticity. If homoscedasticity is satisfied, then the standard error of estimate can be used as an index of the accuracy of prediction for the regression line.

Heteroscedasticity

Sources of Heteroscedasticity

There are four types of problems that can result in heteroscedasticity. They are (a) measurement error, (b) different numbers of cases aggregated at different levels of the predictor variable, (c) systematic variance of the dependent variable at different levels of the predictor variable, and (d) incomplete specification of the regression model. Incomplete specification of the regression model exists when the regression model doesn't include one or more predictor variables that influence the dependent variable.

The first type of problem, measurement error, can result in data that do not accurately reflect the true value of the variables of interest in the population. This can be especially troublesome if the error is systematic. For example, consider an experiment of psychotherapeutic efficacy collecting data at three different locations. If measurements at one site were systematically different than those at the other two sites due to less validity of the measure at one site compared to the other, a pattern of heteroscedasticity could appear in the data.

The second type of problem, different numbers of cases aggregated at different levels of the predictor variable, parallels the phenomenon observed in computing the standard error of the mean. As larger sample sizes are used, the variability of the sampling distribution decreases and becomes more like the normal distribution (Hinkle, Weirsma, & Jurs, 1998). If a different quantity of cases is measured at each level of the DV, the shape of the sample distribution will be different each level. To continue with the psychotherapeutic efficacy example, if the sample included 100 female participants and only 10 male participants and there was no difference in their true depression scores, we would expect more variance in females' depression scores than the males' scores. An atypical depression score for one of the 10 men would affect their group variance more than an atypical depression score would affect group variance for the women. Those differences in variance could result in heteroscedasticity in a regression analysis that grouped the men and women together to predict depression.

The third category of problem, systematic variance in the data, can occur in time series measurements, or may occur as a property of the phenomenon being measured. For example, in the aforementioned study, spontaneous recovery of symptoms might result in reduced differences between groups over time without respect to the predictor variables. The variance between groups would be less at the end of the study than at the beginning, and if there were no variable in the regression equation to explain the difference, the β weights for the regression equation would be inaccurate. See Thorndike (1942) for a discussion of problems caused by preexisting groups that could result in systematic variance in the data.

The fourth category of problem involves incomplete specification of the regression model. Consider the scenario in which some participants in one group of the example efficacy experiment begin taking an antidepressant medication at about the same time they begin the experimental treatment. Since psychopharmacological treatment is known to decrease depression in many cases, it is very likely that an experiment that failed to include the drug treatment in the model would not accurately predict the outcome. If this difference occurred systematically, heteroscedasticity might occur, reducing the accuracy of the linear regression model.

Consequences of Heteroscedasticity

Heteroscedasticity is a problem in regression because the regression line, (or plane or hyperplane), that is created is not the uniformly best-fit line throughout the score distribution. This is due to the fact that the standard error of the dependent variables at each level of the predictor variables creates a different degree of correlation at different levels of the predictor variable. This difference creates a directional bias in the confidence intervals used to calculate the β weights used in the regression equation. If the correlation between the predictor variable (X) and the variance of the dependent variable at some point j ($r_{Xj}\sigma_j^2$) is greater than zero, then the confidence intervals for β will be too narrow and the fit for the line will appear to be statistically significant when it is not. Conversely, if the correlation between the predictor variable and the variance of the dependent variable is less than zero at some point j ($r_{Xj}\sigma_j^2$), then the confidence intervals for β will be too wide and the fit for the line will appear to not be statistically significant when it is.

Barry and Feldman (1985) noted that regression is fairly robust with respect to violations of homoscedasticity, and yet it is possible for some data to violate the assumption so thoroughly that incorrect conclusions will be drawn from the data. The reader may well correctly conclude after considering the classes of problems causing heteroscedasticity that good experimental design is a key factor in preventing problems caused by heteroscedasticity in the data. In addition, the careful researcher will benefit from carefully examining the data to ensure that heteroscedasticity does not result in a regression analysis that fails to accurately reflect the data being analyzed.

Detecting Heteroscedasticity

Heteroscedasticity can be detected through both graphical and mathematical methods. Graphical detection of heteroscedasticity involves visual inspection of a scatterplot of a graph such as a scatterplot of normalized residuals with the dependent variable. The use of graphical means to understand patterns in the data is an important technique for statistical analysis. This is one example of an important means of analysis called Exploratory Data Analysis (Behrens, 1997; Wilkinson, 1999).

Graphical Detection. Dryer and Smith (1998) suggested three scatterplots that can be used for determining the majority of problems of heteroscedasticity. They are (a) residual by time order of data (if known), (b) residual by Yhat values, (c) residual by predictor variable. The stylized scatterplots in Figure 1 illustrate a funnel, band, and curved band. These patterns may be observed in any of the three recommended residual plots. The following discussion of correction will describe possible reasons for each of the three patterns, as well as possible ways to correct the problem to allow for a more accurate regression estimation.

Mathematical Detection. The second approach to detecting heteroscedasticity involves computing test statistics to detect different levels of variance across the range of the predictor variable. Dryer and Smith asserted that test statistics are problematic because it is difficult to know at what level of difference the mathematical difference in residuals is meaningful. A number of tests are available. Many tests only detect a specific pattern of heteroscedasticity, so it is still important to understand the patterns in the data even if mathematical methods are to be used. The Goldfelt-Quandt procedure (Goldfelt & Quandt, 1969) provides a mathematical, test-statistic-based approach to detecting monotonic increases in the residual (error) term as might be seen in the funnel pattern portrayed in Figure 1. In the Goldfelt-Quandt procedure, the data are divided into roughly half, with the middle 10% left out. A regression line is computed for the two groups of data, and the sums of squares of the residuals for the two groups are compared using an F test to determine whether there is a statistically significant difference between the two groups. If there is a statistically significant difference between the two groups, then there is probably heteroscedasticity in the data. Note that this technique would not identify heteroscedasticity of a different type, as would be reflected in a curved or linear band depicted in Figure 1. The interested reader can find additional parametric techniques that can identify other forms of heteroscedasticity such as perturbations in variance *within* the range of the dependent variable in Harrison and McCabe (1979) and Kadiyala and Oberhelman (1984).

Examples

The following examples use the small data set presented in Table 1 to demonstrate various patterns in the residuals. Each example creates a regression equation between the variable X as the predictor variable and one of the Y values (Y0, Y1, Y2, Y3) as the dependent variable. Appendix A includes the SPSS code used for the analysis and to create the graphs.

For the first example, table 2 presents the standard deviation of the dependent variable at each level of the predictor variable to show explicitly the degree of variability within the dependent variable. The first two examples illustrate data that satisfy the assumption of homoscedasticity and further illustrate how uniform differences in variance are not a threat to the assumption. Examples 3 and 4, using dependent variables Y2 and Y3, illustrate different forms of heteroscedasticity. Three graphs are created for each example: a scatterplot of the predictor variable (X) and the dependent variable (Y0, or Y1, ...), a scatterplot of the residuals and the predicted (Yhat) values, and a scatterplot of the residuals and the predictor variable.

Figure 2 is a scatterplot between X and Y0. The regression line touches almost all the points in the scatterplot and that correspondence is reflected in the high R^2 (Rsq in the graph) of .9995. Figure 3 illustrates the graphical method of checking for heteroscedasticity. In this case, the points of the scatterplot are evenly distributed, with no patterned relationship between the residuals and the predictor variable. The values in the Y0 column in Table 2 confirm our conclusion of homoscedasticity, with very similar standard deviations at each level of X.

For the second example, the dependent variable Y1 was created by multiplying each case of Y0 by 500. The application of a multiplicative constant changes the standard deviation by the amount of the multiplier. In this case, the standard deviation of Y0 is 50.81, and the standard deviation of Y1 is equal to 500 times 50.81: 25,403.79. However, as illustrated in Figure 4, the change in standard deviation of the dependent variable does not change the relationship of the variance of two variables. The plot of X with residuals for the Y1 regression equation in Figure 5 confirms that homoscedasticity is unaffected. The column for Y1 in Table 2 confirms our conclusion, with very little difference in standard deviations of Y1 at each level of X.

For the third example, the Y2 dependent variable demonstrates a monotonically increasing deviation of variables (see Figure 6). We know from looking at the examples with Y0 and Y1 that the problem is not the degree of deviation of the dependent variable. This example illustrates the problem is the deviation *across the range* of the predictor variable. Note in Table 2 that the standard deviations for Y2 increase monotonically, confirming the pattern observed in the plot of X with residuals in Figure 7. This pattern might be expected in an experiment in which the model using only X is inadequate to predict the Y value, and including an additional variable might increase the explanatory power of the regression equation substantially beyond that reached using only X as a predictor.

For the fourth example, Y3 dependent variable demonstrates a single perturbation in the level of distribution of the dependent variable as illustrated in Figure 8. Note that, despite the obvious pattern in the data, the R^2 for this equation is only .0008, a useless prediction. The plot of the predictor variable and the residuals in Figure 9 dramatically illustrates the difference in distribution of the dependent variable at the middle point. This is further confirmed in the column for Y3 in Table 2 where the standard deviations of Y for X=1 and X=5 are very different than the standard deviation for X=3. In addition to the possibility of inadequate model specification, it is possible that there may be one or more missing variables that could explain the difference in the observations at that point. This pattern of data could also be cause by a measurement problem for either the predictor variable or the dependent variable or both.

Resolving Problems caused by Heteroscedasticity

Finding heteroscedasticity in the data is a sign to the researcher that she or he should stop and think about the data. Which of the categories of problems is causing heteroscedasticity? The answer to this question will guide the researcher in determining which method, if any, can be used to achieve optimal analysis of the data. Referring back to the three patterns that may indicate heteroscedasticity (funnel, band, or curved band), Draper and Smith (1998) recommended the following corrections that the researcher may consider to correct heteroscedastic data.

The funnel pattern in any of the three residual scatterplots may be corrected by using weighted least squares regression. For example, software packages such as SPSS (SPSS, 1998) provide a Weighted Least Squares (WLS) procedure for creating a weight that, when applied to the measured variable, reduces the importance of cases less likely to be important. Importance is assumed to be reflected in the amount of variance. Those cases that vary more than others are assumed to be less accurate and less important in the regression problem, so they are given a smaller weight than cases that vary less (Weisberg, 1985, p. 82). Stevens (1996, pp. 92-94) gives examples of real data that vary differently at different values of the predictor variable.

The variance-importance assumption of the WLS procedure may not be appropriate for all data. Transforming the dependent variable is an alternative solution to funnel patterns in the Yhat and predictor variable residual scatterplots. Some common transformations include squaring or computing the log of the DV. The researcher would have to make the decision to transform the data based on past research, the nature of the measurements, and his or her values.

Whatever decision is made must consider the meaningfulness of the results after the analysis is complete. A transformation that results in a perfect regression line is not very helpful if there is no reasonable interpretation of the transformed measured variable.

The band pattern in the time order scatterplot may be corrected by adding a first-order term for time. The band pattern in the Yhat scatterplot indicates an error in analysis an incorrect omission of β_0. The band pattern in the X scatterplot indicates an error in calculations, or that the first-order effect of X_j was not removed.

The curved band pattern in the time order scatterplot may suggest a need to add first- and second-order terms for time. The curved band pattern in the other plots may suggest the need to add extra terms to the model or to transform the dependent variable (Y).

It may be that some problems cannot be resolved by other analytical methods or that important variables were not included in the data. The researcher may be forced to use other methods other than regression to analyze the data. Although this might be seen as a failure in the experiment, most would agree that it would be better to abandon a bad study than to publish a meaningless, unreplicable analysis based on faulty assumptions about the data.

Weighted Least Squares Example

The following example demonstrates two methods for estimating a weight for use in a weighted least squares regression. The first method uses the SPSS WLS procedure, the second method is a manual procedure based on an example in Draper and Smith (1998). The manual procedure is presented to give the reader a better idea of what goes on inside the automated SPSS WLS routine. The data, from Draper and Smith, are listed in Table 3. The listing for the SPSS commands is in Appendix B.

Regarding the SPSS automated procedure, Figures 12 and 13 show the funnel shaped scatterplots for the ordinary least squares residuals with Yhat values and the predictor variable. [The recommended time variable plot was not created because there was no time of observation in this data set.] These plots suggest heteroscedasticity; for the purposes of the example, assume that the problem domain suggests the best course of means of analysis is weighted least squares regression. The SPSS procedure for performing WLS is described in the following steps:

1) Use the WLS routine to create a weight variable (wgt_1), in this example we instructed SPSS to try weights from -2 to 2 in steps of 0.5.
2) Perform a weighted regression using the weight created in the previous step by adding the REGWGT=wgt_1 parameter to the REGRESSION command. Since SPSS won't automatically create residual plots when the REGWGT parameter is included, the residual and Yhat values are saved as new variables.
3) Manually create scatterplots for the Predictor variable and Yhat with residuals using the values saved in the last step, (see Figures 14 and 15).

A manual process for estimating the weighted least squares regression is described by Draper and Smith (1998). We can be fairly confident that the two procedures are the same because they result in very similar effect sizes and scatterplots. The adjusted effect size (R^2) of the two WLS routines are .919 for the manual procedure and .947 for the automated procedure. The plots (Figures 14 and 15 for the automated process and 16 and 17 for the manual process) are almost identical. Appendix B also includes the SPSS commands to perform this manual procedure. There are four primary steps ion the procedure.

1) Average the values of the predictor variable that are approximately the same (e.g., 10.21, 10.22, 10.30 are all grouped together) and compute the variance for the

corresponding group of DVs (see Table 4). The plot of these indicates a quadratic relationship, so add a squared form of the mean of the PV group. The SPSS WLS routine performs this procedure by trial and error, looking for the best fir from the start point to the end point in steps of the size specified in the WLS command.

2) Estimate the least squares regression equation for the grouped means and variances. The data for this equation comes from the summarized means and variances in Table 4. The resulting regression equation is

$$s_{ej}^2 = 1.533 - 0.733 * \text{mean}(X) + 0.826 * \text{mean}(X)^2.$$

3) Create the weight variable using $1/ s_{ej}^2$.

4) Use the same procedure as in the other example to perform the WLS routine.

A comparison of scatterplots for Yhat and PV for the SPSS WLS routine (Figures 14 and 15) or for the manual WLS routine (Figures 16 and 17) reveals mostly evenly distributed residuals. Compare this to the scatterplots for the OLS Yhat and PV (Figures 12 and 13), showing a pattern of increasing variance as the predictor variable (X) gets larger.

Summary

This paper has described the concept of homoscedasticity, an important assumption of linear regression. Homoscedasticity is consistency in variance of the dependent variable across the range of values for the predictor variable. Different causes of heteroscedasticity were described, along with graphical and mathematical methods for detecting heteroscedasticity in the data. Possible solutions to problems of heteroscedasticity were briefly mentioned.

Although statistical assumptions are usually included as one of several threats to experimental validity (Campbell & Stanley, 1963; Heppner, Kivlighan, & Wampold, 1999), the prior examples demonstrate that all four types of validity: statistical, internal, construct, and external, influence one another to contribute to or detract from investigation soundness. Elements of good internal validity such as careful measurement practice can contribute to statistical validity. The third example illustrated how threats to external validity such as interaction of selection and treatment can also result in heteroscedasticity. As the fourth example in the hypothetical psychotherapy study illustrated, threats to internal validity such as experimental design based on an inadequate construct can result in heteroscedasticity. Checking for the assumption of homoscedasticity can improve the experimenter's confidence in the internal and external validity of the experiment.

The importance of ensuring that required assumptions for a statistical procedure are met was illustrated in the examples. Using linear regression on data that violates important assumptions can result in incorrect conclusions from the design. Good statistical validity goes hand-in-hand with the other aspects of experimental design. Statistical validity can be improved by ensuring that internal and external experimental validity (Campbell & Stanley, 1963) are carefully considered prior to collecting the data. At the same time, checking statistical assumptions such as homoscedasticity can improve the confidence we have in our results and interpretations.

References

Behrens, J.T. (1997). Principles and procedures of exploratory data analysis. Psychological Methods, 2(2), 131-160.

Berry, W.D., & Feldman, S. (1985). Multiple regression in practice. Beverly Hills, CA: Sage Publications.

Campbell, D.T., & Stanley, J.C. (1963). Experimental and quasi-experimental designs for research. Chicago: Rand McNally.

Draper, N. R., & Smith, H. (1998). Applied Regression Analysis. New York: John Wiley & Sons.

Guilford, J.P., & Fruchter, B. (1978). Fundamental Statistics in Psychology and Education. New York: McGraw-Hill.

Harrison, M.J., & McCabe, B.P.M. (1979). A test for heteroscedasticity based on ordinary least squares residuals. Journal of the American Statistical Association, 74, 494-499.

Heppner, P.P., Kivlighan, Jr., D.M, Wampold, B.E. (1999). Research Design in Counseling (2nd ed.). Belmont, CA: Wadsworth.

Hinkle, D.E., Wiersma, W., & Jurs, S.G. (1998). Applied Statistics for the Behavioral Sciences. Boston: Houghton Mifflin.

Kadiyala, K.R., & Oberhelman, H.D. (1984). Alternative tests for heteroscedasticity of disturbances: A comparative study. Communications in Statistics: Theory and Methods, 13, 987-1001.

Pedhazur, E.J. (1982). Multiple Regression in Behavioral Research: Explanation and Prediction. New York: Holt, Rinehart, and Winston.

Snyder, P., & Lawson, S. (1993). Evaluating results using corrected and uncorrected effect size estimates. Journal of Experimental Education, 61(4), 334-349.

SPSS, Inc. (1998). SPSS Version 7.5. Chicago: Author.

Stevens, J. (1996). Applied Multivariate Statistics for the Social Sciences. Mahwah, N.J.: Lawrence Erlbaum Associates.

Thompson, B. (1992, April). Interpreting regression results: beta weights and structure coefficients are both important. Paper presented at the annual meeting of the American Educational Research Association. San Francisco. (ERIC Document Reproduction Service No. ED 344 897)

Thompson, B. (1998, April). Five methodology errors in educational research: The pantheon of statistical significance and other faux pas. Invited address presented at the annual meeting of the American Educational Research Association, San Diego. (ERIC Document Reproduction Service No. ED 419 023)

Thompson, B. (1999, April). Common methodology mistakes in educational research, revisited, along with a primer on both effect sizes and the bootstrap. Invited address presented at the annual meeting of the American Educational Research Association, Montreal. (ERIC Document Reproduction Service No. 429 110)

Thompson, B., & Borrello, G.M. (1985). The importance of structure coefficients in regression research. Educational and Psychological Measurement, 45, 203-209.

Thorndike, R.L. (1942). Regression fallacies in the matched groups experiment. Psychometrika, 7(2), 85-102.

Weisberg, S. (1985). Applied Linear Regression. New York: John Wiley and Sons.

Wilkinson, L. (1999). Graphs for research in counseling psychology. The Counseling Psychologist, 27, 384-407.

Wilkinson, L., & The APA Task Force on Statistical Inference. (1999). Statistical methods in psychology journals: Guidelines and explanations. American Psychologist, 54, 594-604. [reprint available through the APA Home Page: **http://www.apa.org/journals/amp/amp548594.html**].

Wohlke, P.L., Elmore, P.B., & Spearing, D.L. (1990). Testing assumptions in multiple regression: Comparison of procedures available in SAS and SPSS-X. Multiple Linear Regression Viewpoints, 17(2), 48-66.

Table 1

Example Data

PV	Y0	Y1	Y2	Y3
1	30	1500	30	30
1	32	1600	90	90
1	31	1550	75	150
1	29	1450	75	210
1	29	1450	75	270
3	90	4500	30	30
3	92	4600	90	35
3	91	4550	150	45
3	90	4500	210	50
3	92	4600	135	55
5	150	7500	30	60
5	152	7600	90	90
5	149	7450	150	150
5	151	7550	210	210
5	150	7500	270	270

Table 2
Standard deviations for Dependent Variables at Different Levels of Predictor Variable

	SD_{Y0}	SD_{Y1}	SD_{Y2}	SD_{Y3}
X=1	1.30	651.92	22.75	94.87
X=3	1.03	651.92	67.42	10.37
X=5	1.14	570.09	94.87	58.91

Table 3

Data from Draper and Smith (1998, p. 226)

X	Y	W
1.15	0.99	1.240280
1.90	0.98	2.182240
3.00	2.60	7.849300
3.00	2.67	7.849300
3.00	2.66	7.849300
3.00	2.78	7.849300
3.00	2.80	7.849300
5.34	5.92	7.436520
5.38	5.35	6.993090
5.40	4.33	6.785740
5.40	4.89	6.785740
5.45	5.21	6.305140
7.70	7.68	0.892040
7.80	9.81	0.844200
7.81	6.52	0.839630
7.85	9.71	0.821710
7.87	9.82	0.812960
7.91	9.81	0.795880
7.94	8.50	0.783420
9.03	9.47	0.473850
9.07	11.45	0.466210
9.11	12.14	0.458480
9.14	11.50	0.453270
9.16	10.65	0.449680
9.37	10.64	0.414350
10.17	9.78	0.311820
10.18	12.39	0.310790
10.22	11.03	0.306720
10.22	8.00	0.306720
10.22	11.90	0.306720
10.18	8.68	0.310790
10.50	7.25	0.280330
10.23	13.46	0.305710
10.03	10.19	0.326800
10.23	9.93	0.305710

Table 4

Means for PV and Standard Deviations for Residuals

Mean(X_j)	Mean$(X_j)^2$	Standard Deviation(e_j)
3.0	9.00	0.0072
5.4	29.16	0.3440
7.8	60.84	1.7404
9.1	82.81	0.8683
10.2	104.04	3.8964

Figure 1. Stylized examples of patterns of heteroscedasticity.

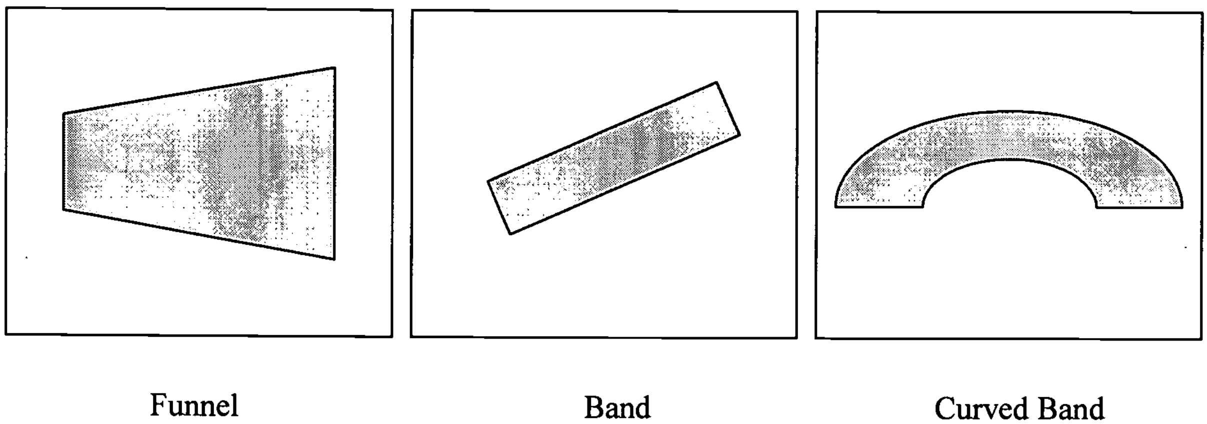


Figure 2. Scatterplot of Small Normal Variance DV (Y0) with PV.

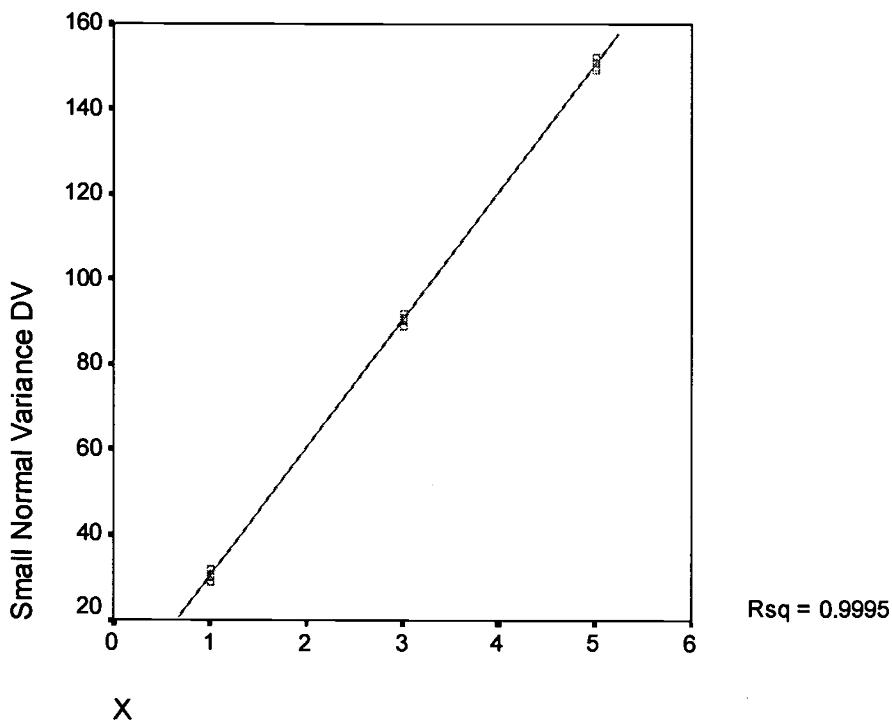


Figure 3. Scatterplot of Small Normal Variance DV (Y0) Residual.

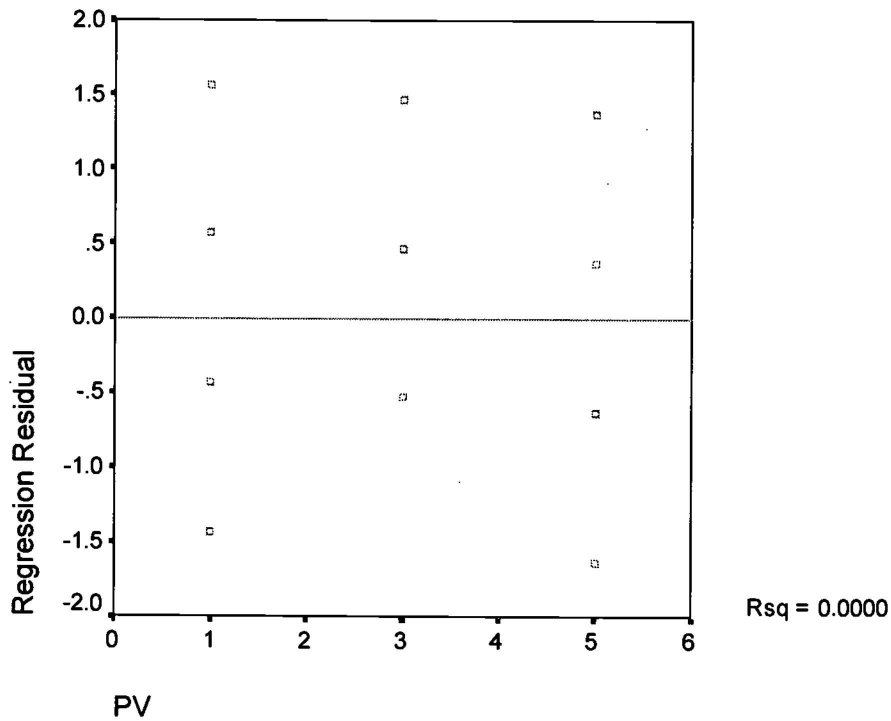


Figure 4. Scatterplot of Large Normal Variance DV (Y1) with PV.

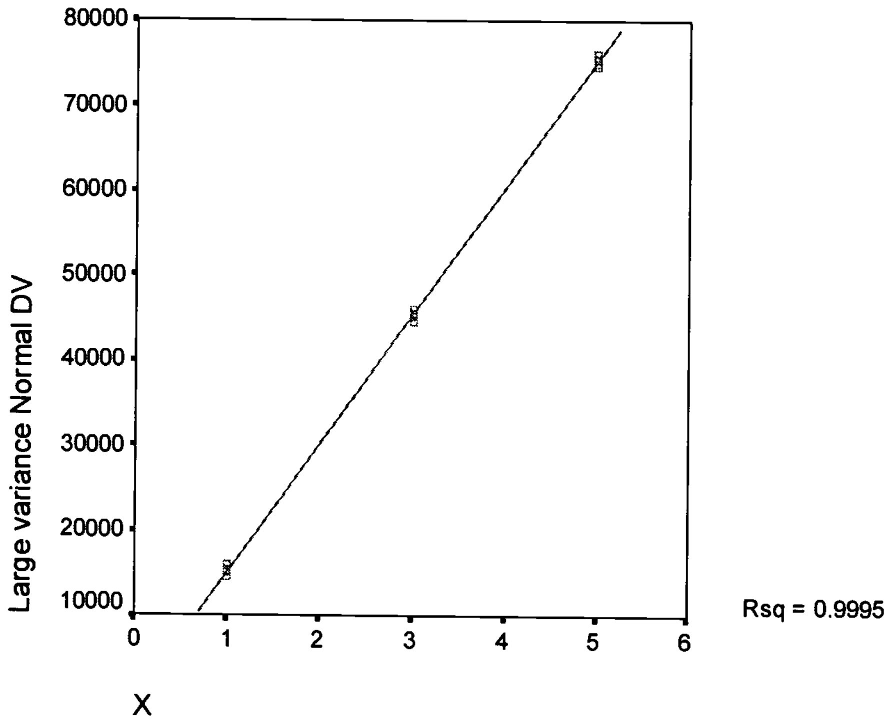


Figure 5. Scatterplot of Large Normal Variance DV (Y1) with Residual.

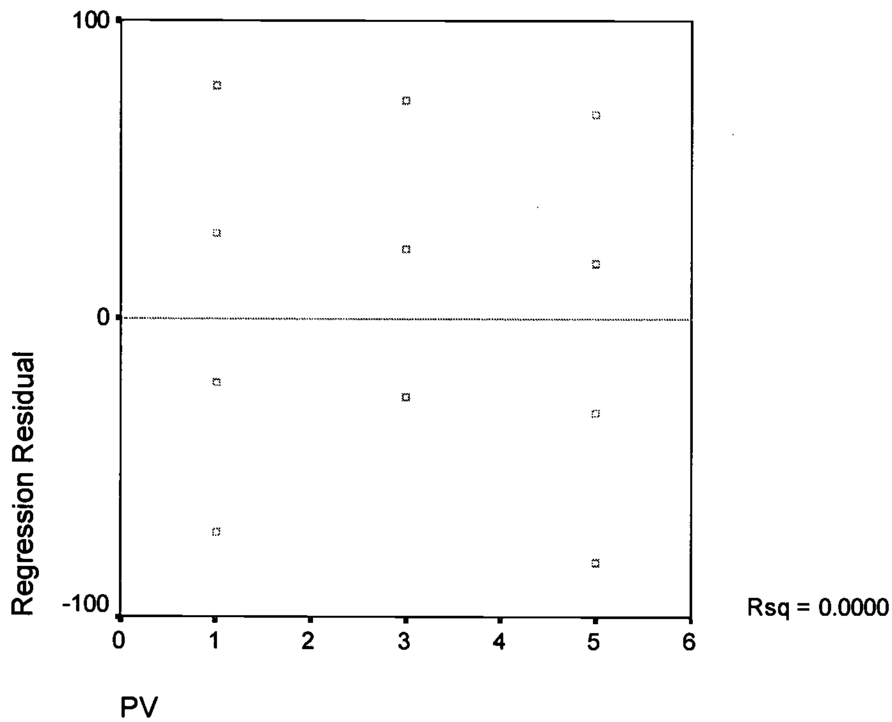


Figure 6. Scatterplot of Skewed DV (Y2) with PV.

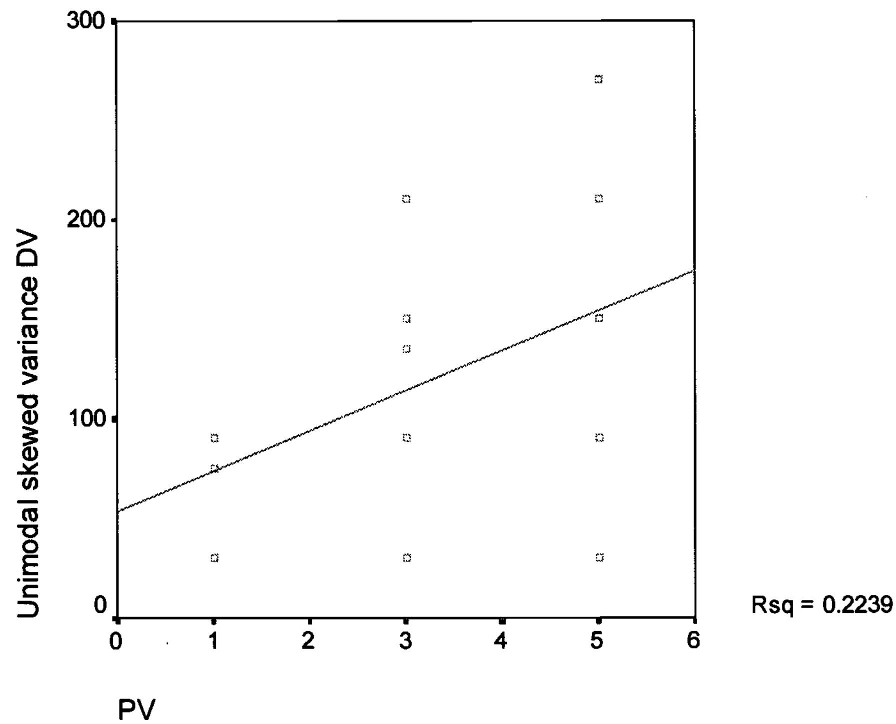


Figure 7. Scatterplot of Skewed DV (Y2) with Residual.

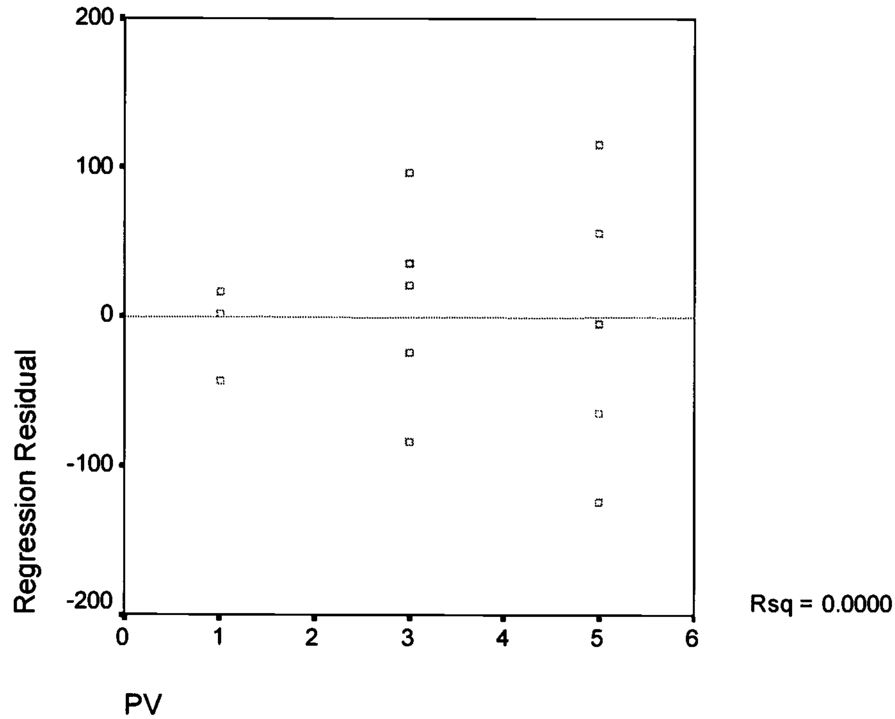


Figure 8. Scatterplot of Bimodally Skewed DV (Y3) with PV.

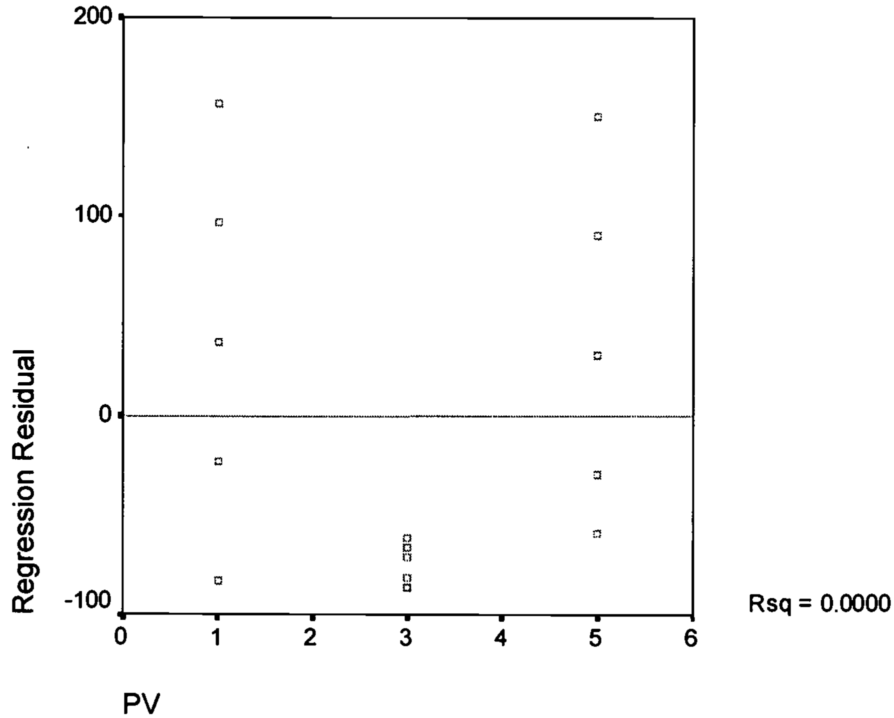


Figure 9. Scatterplot of Bimodally Skewed DV (Y3) with Residual.

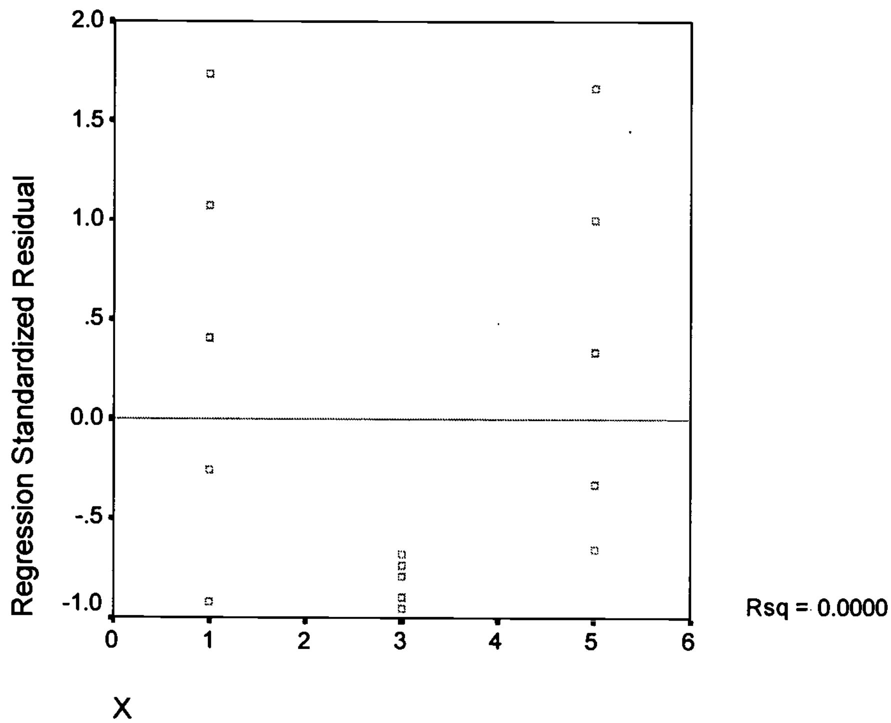


Figure 10. Normal Probability (P-P) Plot of Standardized Residuals for Y0.

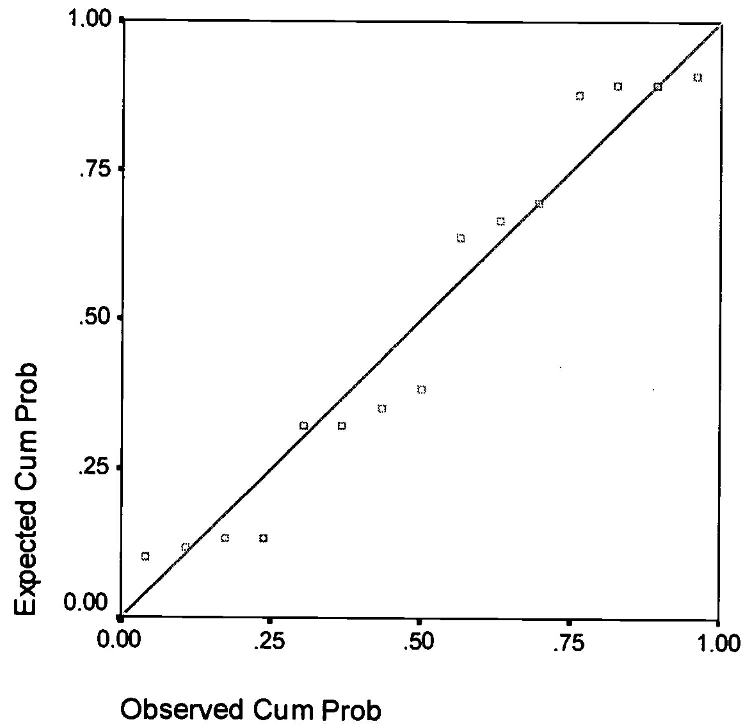


Figure 11. Normal Probability (P-P) Plot of Standardized Residuals for Y3.

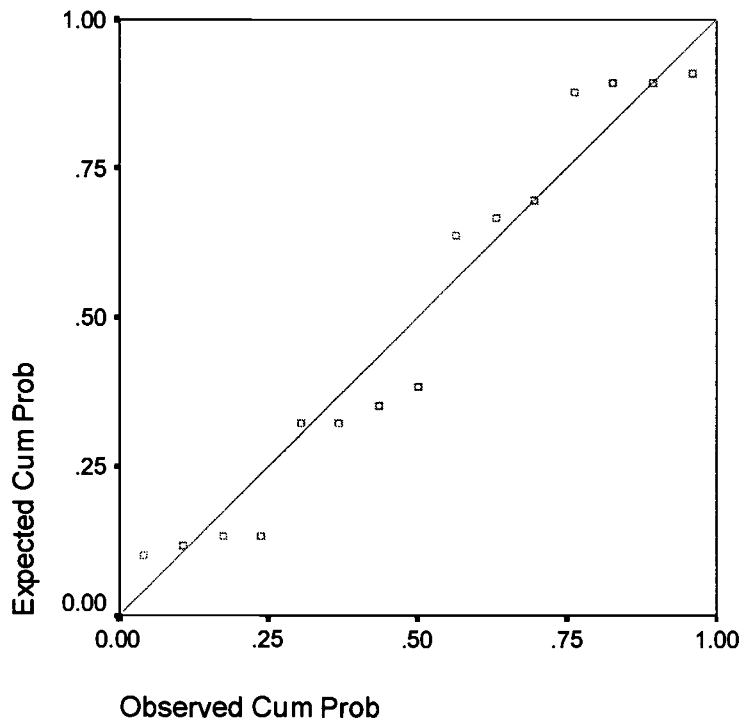


Figure 12. Scatterplot of X with OLS Residual.

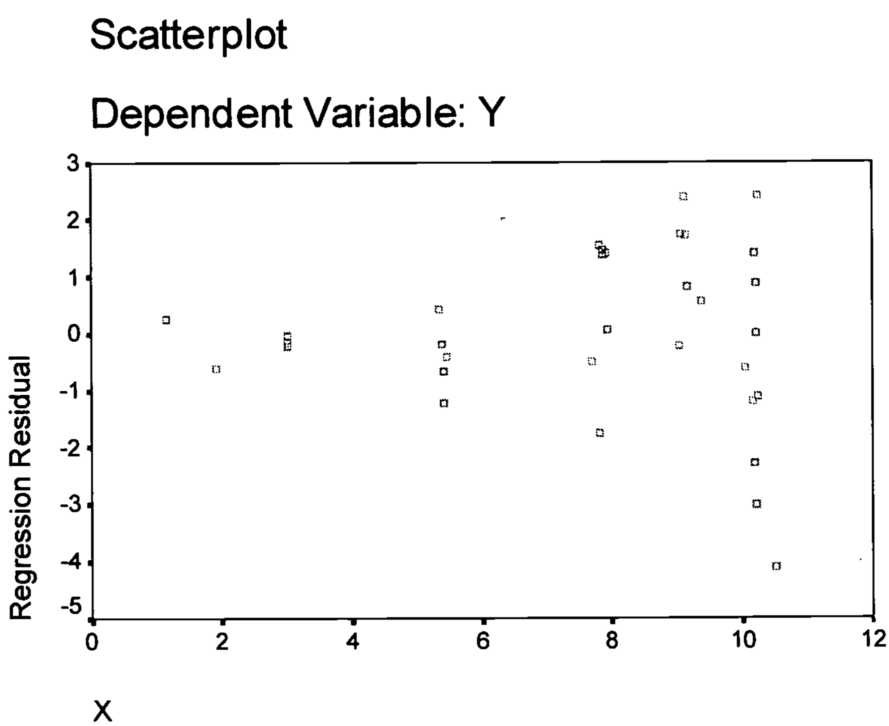


Figure 13. Scatterplot of Yhat with OLS Residual.

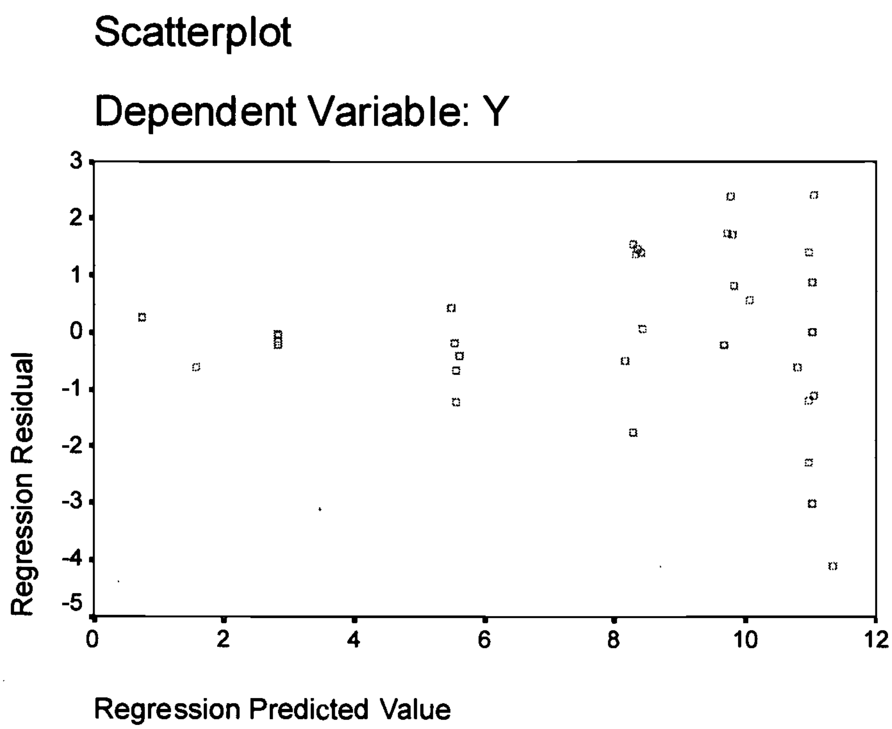


Figure 14. Scatterplot of w With SPSS-Generated Weighted Least Squares Residual.

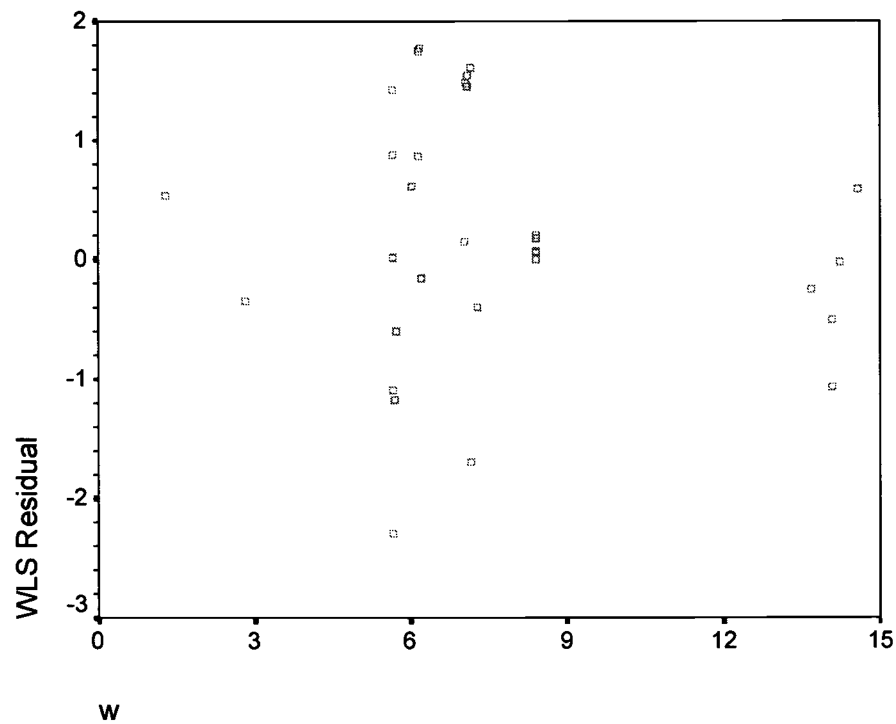


Figure 15. Scatterplot of Yhat With SPSS Generated Weighted Least Squares Residual.

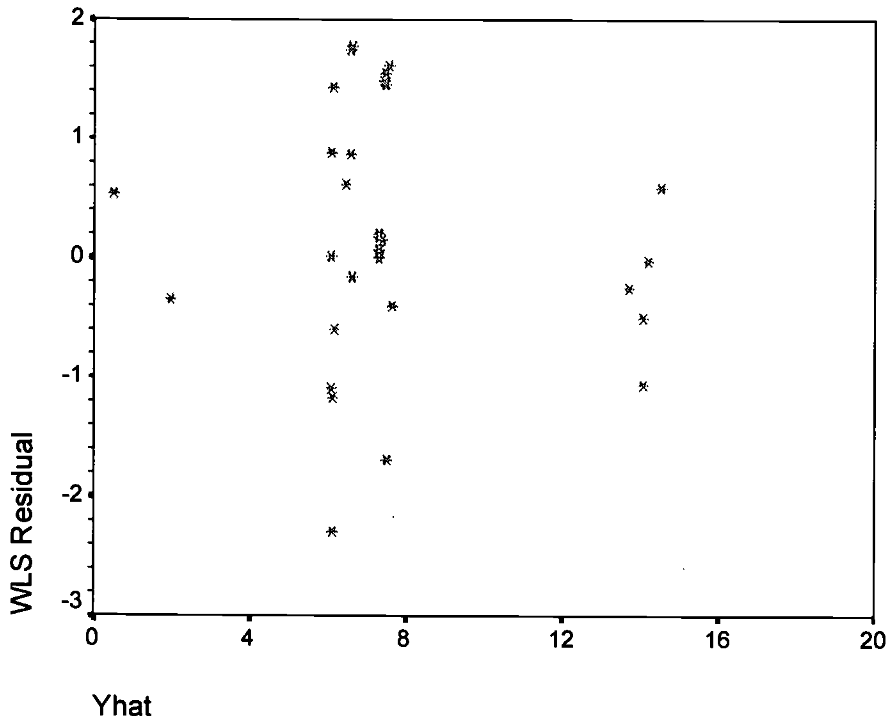


Figure 16. Scatterplot of w With Manually Generated Weighted Least Squares Residual.

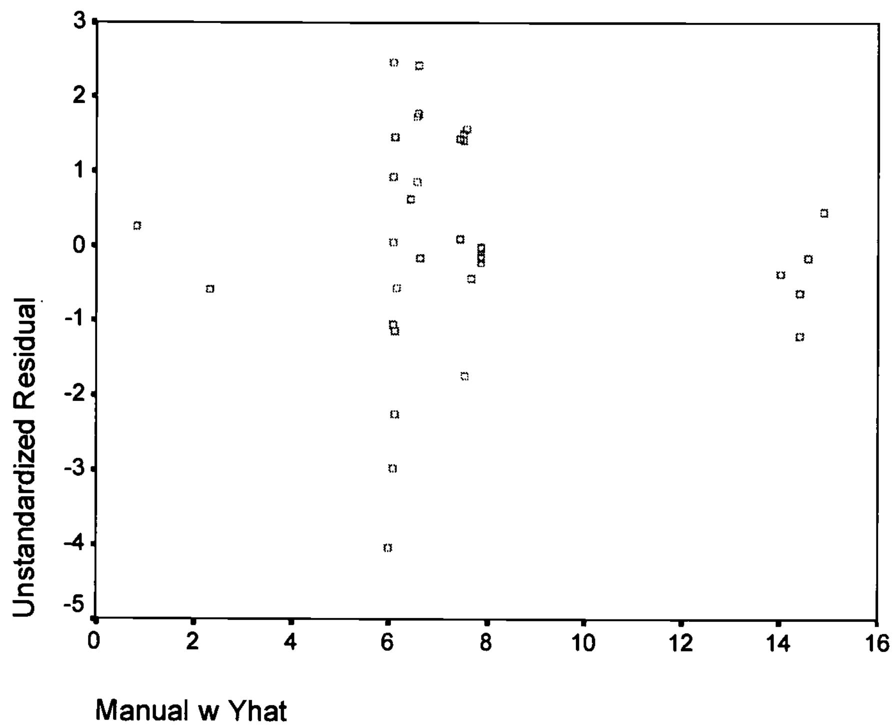
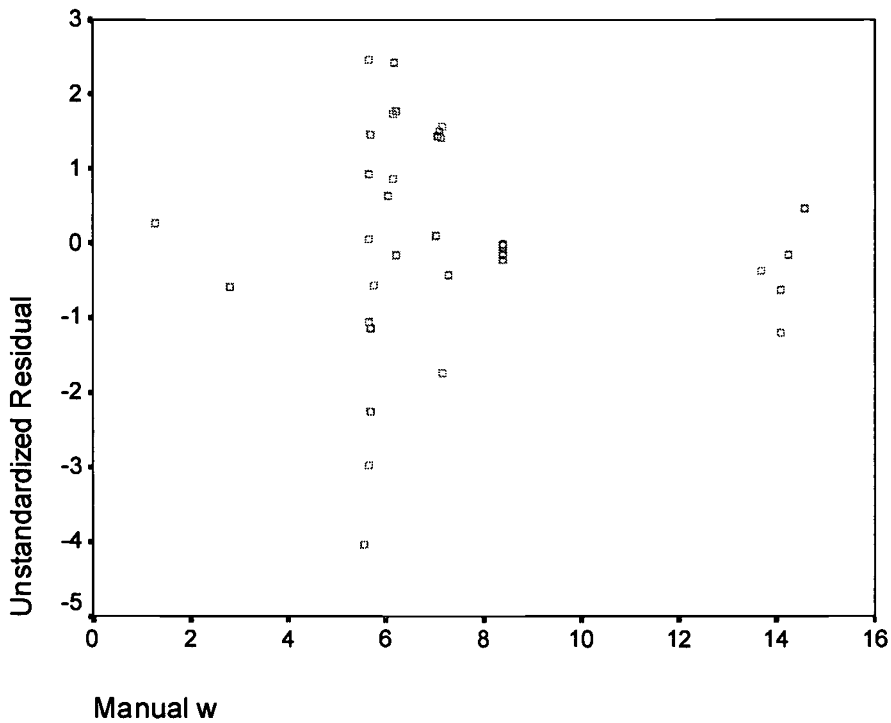


Figure 17. Scatterplot of Yhat With Manually Generated Weighted Least Squares Residual.



Appendix A

Listing for Residual Plot Examples

```
REGRESSION
 /MISSING LISTWISE
 /STATISTICS COEFF OUTS R ANOVA
 /CRITERIA=PIN(.05) POUT(.10)
 /NOORIGIN
 /DEPENDENT y0
 /METHOD=ENTER X
 /SCATTERPLOT=(y0, X)  (*ZRESID, X)
/RESIDUALS NORM(ZRESID)  .

REGRESSION
 /MISSING LISTWISE
 /STATISTICS COEFF OUTS R ANOVA
 /CRITERIA=PIN(.05) POUT(.10)
 /NOORIGIN
 /DEPENDENT y1
 /METHOD=ENTER X
 /SCATTERPLOT=(y1, X)  (*ZRESID, X)
/RESIDUALS NORM(ZRESID)  .

REGRESSION
 /MISSING LISTWISE
 /STATISTICS COEFF OUTS R ANOVA
 /CRITERIA=PIN(.05) POUT(.10)
 /NOORIGIN
 /DEPENDENT y2
 /METHOD=ENTER X
 /SCATTERPLOT=(y2, X )  (*ZRESID, X)
/RESIDUALS NORM(ZRESID) .

REGRESSION
 /MISSING LISTWISE
 /STATISTICS COEFF OUTS R ANOVA
 /CRITERIA=PIN(.05) POUT(.10)
 /NOORIGIN
 /DEPENDENT y3
 /METHOD=ENTER X
 /SCATTERPLOT=(y3, X )  (*ZRESID, X)
/RESIDUALS NORM(ZRESID)  .
```

Appendix B
Listing for Correction Examples

```
REGRESSION
 /MISSING LISTWISE
 /STATISTICS COEFF OUTS R ANOVA
 /CRITERIA=PIN(.05) POUT(.10)
 /NOORIGIN
 /DEPENDENT y
 /METHOD=ENTER X
 /SCATTERPLOT=(y, X)   (*RESID, X) (*RESID ,*PRED )
 /RESIDUALS NORM(ZRESID) .

TITLE 'Default spss way to do weight estimation' .
* Weight Estimation.
WLS y  WITH x
 /SOURCE x
 /POWER -2 TO 2 BY 0.5
 /CONSTANT
 /SAVE WEIGHT
 /PRINT ALL.

REGRESSION
 /MISSING LISTWISE
 /REGWGT=wgt_1
 /STATISTICS COEFF OUTS R ANOVA
 /CRITERIA=PIN(.05) POUT(.10)
 /NOORIGIN
 /DEPENDENT y
 /METHOD=ENTER x
 /SAVE PRED RESID .

COMPUTE sqrtwyh = sqrtw *  pre_1 .
EXECUTE .

COMPUTE sqrtwx = sqrtw * x .
EXECUTE .

GRAPH
 /SCATTERPLOT(BIVAR)=sqrtwyh WITH res_1
 /MISSING=LISTWISE .

GRAPH
 /SCATTERPLOT(BIVAR)=sqrtwx WITH res_1
 /MISSING=LISTWISE .

title 'Use the Draper & Smith (1998) process to estimate the weight' .
subtitle 'This yields essentially the same result as the WLS routine above'.

title 'Find a quadratic least squares relationship for X and variance of Y' .
subtitle 'For each set of repeats or near repeats of X' /
REGRESSION
 /MISSING LISTWISE
 /STATISTICS COEFF OUTS R ANOVA
 /CRITERIA=PIN(.05) POUT(.10)
```

```
 /NOORIGIN
 /DEPENDENT d_s_ssq
 /METHOD=ENTER d_s_x d_s_xsq  .

title 'Use the weights from the above relationship to compute an estimated weight matrix w' .
subtitle ' for the weighted least squares regression using the meanx and meanxsq values'.
COMPUTE w = 1/(1.5329  - (.7334 * x) + ( .0883 * x**2)) .
EXECUTE .

title 'Use the spss weighted regression procedure using w as the weight'.

REGRESSION
 /MISSING LISTWISE
 /REGWGT=w
 /STATISTICS COEFF OUTS R ANOVA
 /CRITERIA=PIN(.05) POUT(.10)
 /NOORIGIN
 /DEPENDENT y
 /METHOD=ENTER x
 /SAVE PRED RESID .

COMPUTE sqrtwyh = sqrtw *  pre_2 .
EXECUTE .

COMPUTE sqrtwx = sqrtw * x .
EXECUTE .

GRAPH
 /SCATTERPLOT(BIVAR)=sqrtwyh WITH res_2
 /MISSING=LISTWISE .

GRAPH
 /SCATTERPLOT(BIVAR)=sqrtwx WITH res_2
 /MISSING=LISTWISE .
```

TM031904

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement (OERI)
Educational Resources Information Center (ERIC)

REPRODUCTION RELEASE
(Specific Document)

I. DOCUMENT IDENTIFICATION:

Title: An Introduction to Graphical & Mathematical Methods for Detecting Heteroscedasticity in Linear Regression	
Author(s): Russel L. Thompson	
Corporate Source:	Publication Date: 1/99

II. REPRODUCTION RELEASE:

In order to disseminate as widely as possible timely and significant materials of interest to the educational community, documents announced in the monthly abstract journal of the ERIC system, *Resources in Education* (RIE), are usually made available to users in microfiche, reproduced paper copy, and electronic/optical media, and sold through the ERIC Document Reproduction Service (EDRS) or other ERIC vendors. Credit is given to the source of each document, and, if reproduction release is granted, one of the following notices is affixed to the document.

If permission is granted to reproduce the identified document, please CHECK ONE of the following options and sign the release below.

[X] Sample sticker to be affixed to document

Check here
Permitting microfiche (4"x 6" film), paper copy, electronic, and optical media reproduction

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

Level 1

Sample sticker to be affixed to document [ ]

or here
Permitting reproduction in other than paper copy.

"PERMISSION TO REPRODUCE THIS MATERIAL IN OTHER THAN PAPER COPY HAS BEEN GRANTED BY

Sample

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

Level 2

Sign Here, Please

Documents will be processed as indicated provided reproduction quality permits. If permission to reproduce is granted, but neither box is checked, documents will be processed at Level 1.

"I hereby grant to the Educational Resources Information Center (ERIC) nonexclusive permission to reproduce this document as indicated above. Reproduction from the ERIC microfiche or electronic/optical media by persons other than ERIC employees and its system contractors requires permission from the copyright holder. Exception is made for non-profit reproduction by libraries and other service agencies to satisfy information needs of educators in response to discrete inquiries."

Signature: [signature]	Position: RES ASSOCIATE
Printed Name: Russell L. Thompson	Organization: TEXAS A&M UNIVERSITY
Address: TAMU DEPT EDUC PSYC COLLEGE STATION, TX 77843-4225	Telephone Number: (409) 845-1831
	Date:

OVER