An Empirical Comparison of Regression Analysis Strategies with Discrete Ordinal Variables

Jeffrey D. Kromrey **Gianna Rendina-Gobioff**
University of South Florida

The Type I error control and statistical power of three tests of regression models incorporating discrete ordinal variables were compared in a Monte Carlo study. Samples were generated from populations measured on discrete ordinal variables representing 5-point and 7-point response scales. Each sample was analyzed using ordinary least squares regression, ordinal multiple regression and cumulative logit models. Factors examined in the Monte Carlo study were the population effect size, number of regressor variables, level of regressor intercorrelation, population distribution shape and sample size. Results suggest that the logistic regression approach evidenced poor Type I error control with small samples or with large numbers of regressors. In contrast, both the ordinary least squares approach and the ordinal multiple regression approach evidenced good Type I error control across the majority of conditions examined. Further, the power differences between these approaches were negligible.

A common response format for the measurement of many variables in the social sciences is a forced-choice, ordinal scale. For example, data may be collected using a Likert scale in which respondents indicate the extent of their agreement or disagreement with a stimulus. Similarly, ordinal rating scales may be used to obtain ratings by participants regarding their perceptions of the frequency or intensity of a target phenomenon.

The level of measurement of this type of variable and the appropriate statistical techniques for the analysis of such data have been the subject of debate for many years. Stevens (1951) defined four levels of measurement (nominal, ordinal, interval and ratio) and the appropriateness of their use within statistical analyses. The subsequently published literature presents conflicting views about the use of parametric statistics with such discrete ordinal variables. Critics of Stevens have proposed that ordinal data can be treated as interval data when parametric statistics are employed (e.g., Borgatta & Bohrnstedt, 1980; Gaito, 1980). In contrast, some researchers have proposed alternative analyses that are purported to more appropriately represent the discrete ordinal nature of such data (Agresti & Finlay, 1997; Cliff, 1994, 1996; Long, 1999), while others have recommended a rescaling of such data to provide a closer approximation to interval-level measurement (Harwell & Gatti, 2001). The long-standing debate about how to treat ordinal variables in statistical analyses is fueled by the need within the social sciences to accurately answer research questions in which measures are obtained using a discrete ordinal response scale.

The purpose of this study was to investigate available analysis options for ordinal level data in the context of multiple regression analysis, and to empirically compare the performance of these analysis options. Multiple regression was the focus of this study because it has proven to be a useful, general purpose tool for data analysis encompassing the analysis of variance and analysis of covariance as special cases. As stated above, the literature reflects disagreement about appropriate analyses of discrete ordinal data. Therefore, one purpose of our work is to raise researchers' awareness of multiple regression analysis options with discrete ordinal variables. Further, the literature base is lacking empirical studies that investigate factors of the research context (such as effect size, number of regressors, sample size, and level of measurement) that may affect the performance of the analysis options. Therefore, a second purpose was to compare the performance of three analysis options with various manipulations of factors under controlled conditions.

Analysis Options with Discrete Ordinal Variables

Through the years, ordinal data have been analyzed using a variety of models, including ordinary least squares regression, variants of logistic regression analysis, and techniques developed specifically for ordinal-level analyses. Each of these methods will be briefly described.

Ordinary Least Squares. Ordinary least squares regression (OLS) analysis essentially ignores the discreteness and the ordinal level of measurement of variables, treating all values as if they represented a continuous interval-level measure. The well-known general linear model in matrix form is given by

$$\boldsymbol{y} = \boldsymbol{X}\boldsymbol{b} + \boldsymbol{\varepsilon} \ ,$$

where $\boldsymbol{y}$ is an n x 1 vector of values for the dependent variable,
$\boldsymbol{X}$ is an n x $(k + 1)$ matrix of observations on the k regressor variables augmented with a unit vector to provide an intercept,
$\boldsymbol{b}$ is a $(k + 1)$ x1 vector of regression coefficients, and
$\boldsymbol{\varepsilon}$ is an n x 1 vector of residuals.

The OLS regression coefficients are obtained as:

$$\hat{\boldsymbol{b}} = (\boldsymbol{X}^T\boldsymbol{X})^{-1}\boldsymbol{X}^T\boldsymbol{y} \ .$$

Hypothesis tests associated with OLS regression are typically of the form $H_0 : \beta_i = 0$, where β_i is the parameter associated with one of the elements of the vector $\boldsymbol{b}$. This test is conducted by obtaining the ratio of the parameter estimate to its standard error and comparing that ratio to a t distribution with df = n – k – 1. A simultaneous test that all of the β_i parameters are equal to zero (equivalent to the test that the population squared multiple correlation coefficient is equal to zero) is obtained by calculating the ratio

$$F = \frac{R^2 / k}{(1 - R^2) / (n - k - 1)}$$

and comparing the result to an F distribution with df = k, n – k – 1.

<u>*Ordinal Multiple Regression*</u>. Cliff (1994) presented an ordinal multiple regression strategy (OMR) and Long (1999) provided a method for estimating confidence intervals around the parameters of this model. The OMR model was developed to predict the ordinal information (that which Cliff refers to as "dominance") in a criterion variable, based upon the observed ordinal information among the predictor variables. That is, OMR predicts the dominance scores on the criterion variable (d_{ihy}) by weighting the observed dominance scores on the regressor variables (d_{ihx}). The dominance scores are given as

$$d_{ihx} = sign\left(x_i - x_h\right),$$

taking the value of 1 if the i^{th} observation has a higher "score" on X than does the h^{th} observation, a value of –1 if the score on X is lower, and a value of zero if the two observations are tied on X. The weights for the OMR model are obtained as

$$\boldsymbol{w} = \boldsymbol{T}_{\boldsymbol{x}}^{-1}\boldsymbol{t}_{\boldsymbol{y}}$$

where $\boldsymbol{w}$ is a k X 1 vector of regression weights,
$\boldsymbol{T_x}$ is a k X k matrix of Kendall tau-a correlation coefficients among the regressors, and
$\boldsymbol{t_y}$ is a k X 1 vector of tau-a correlations between the regressors and the criterion variable.

Confidence intervals and hypothesis tests about the OMR weights ($\boldsymbol{w}$) are conducted using the standard normal (z) distribution. That is, the confidence interval for the j^{th} weight, described by Long (1999) is $w_j \pm Z_{(1-\alpha/2)}\hat{\sigma}_{wj}$ and a test of the null hypothesis that the population weight is zero is conducted by dividing the sample weight by its standard error and comparing the ratio to a critical value of Z.

<u>*Logistic Regression*</u>. A cumulative logit model for ordinal criterion variables (LR) was described by Agresti and Finlay (1997). The LR analysis models the probability of a value on the dependent variable being in the j^{th} category or lower, that is, modeling $P(y \le j)$ based on the values of predictor variables. In logit form, the model is

$$\log\left(\frac{P(y \le j)}{P(y > j)}\right) = \alpha_j + \beta X$$

where α_j is the intercept for the j^{th} category of the ordinal response, and
β is a k X 1 vector of regression weights for the predictors.

Parameter estimates for the LR models are usually obtained using maximum likelihood methods. Inference in the cumulative logit model is conducted using Wald tests for the individual regression parameters (ratios of the estimates to their standard errors), and likelihood ratio tests to obtain a simultaneous test that all parameters are equal to zero.

Method

This research was a Monte Carlo study in which random samples were generated under known and controlled population conditions. In the Monte Carlo study, samples were generated from populations of discrete ordinal variables. In each sample the data were analyzed using the OLS, OMR, and cumulative logit regression approaches.

The Monte Carlo study included six factors in the design. These factors were (a) the true population effect size of the individual regressors (populations were simulated with effect sizes corresponding to small, medium and large values of Cohen's effect size f^2, as well as a null condition (effect size of zero), (b) number of regressor variables (with k = 2, 5 and 10 regessors), (c) correlation between regressor variables (with r_{12} = .10 and .30), (d) sample sizes (with n = 5*k, 10*k and 100*k), (e) level of measurement of variables (discrete, ordinal variables were investigated as 5-option response scales, and 7-option response scales), and (f) population distribution shape (conditions were simulated in which the discrete ordinal variables were uniform, unimodal symmetric, and unimodal skewed).

Generation of Pseudo-populations. Because the population parameters corresponding to regression weights in the models being investigated for discrete ordinal data cannot be determined analytically, the simulation study was conducted by generating a pseudo-population for each condition, then drawing random samples with replacement from these pseudo-populations. The pseudo-populations consisted of 10,000 observations randomly generated from the corresponding population (level of discreteness of the measurements, distribution shape and degree of relationship between the regressors and the criterion variable).

The data were generated by transforming uniform random variates obtained from the RANUNI function in SAS, using a modification of the technique described by Bradley and Fleisher (1994), and operationalized by Ferron, Yi, and Kromrey (1997). In this method, a population correlation matrix, R, based on discrete ordinal variables is constructed by an iterative process in which large simulated samples (n = 100,000) are generated from an approximation of R, $(\tilde{R})$. The observed correlation matrix obtained from this large sample $(\hat{R})$ is compared elementwise to R, and the residuals $(R-\hat{R})$ are used to adjust the generating matrix $\tilde{R}$. This sequence of large sample generation, matrix estimation, and adjustment of $\tilde{R}$ continues until the process converges. The resulting matrix, $\tilde{R}$, is used as a template to generate correlated discrete data for the Monte Carlo study.

Each analysis model was calculated on all observations in these pseudo-populations providing values that served as the corresponding population parameters. The Monte Carlo simulation was then conducted by randomly sampling observations, with replacement, from these pseudo-populations.

The research was conducted using SAS/IML version 8.1. Conditions for the study were run under Windows 98. Discrete random variables were generated using the RANUNI function of SAS. A different seed value was used in each execution of the program and the program code was verified by hand-checking results from benchmark datasets.

For each condition investigated, 5,000 samples were generated. The use of 5,000 samples provides adequate precision for the investigation of the sampling behavior of these statistics. For example, 5,000 samples provides a maximum 95% confidence interval width around an observed proportion that is $\pm$ .014 (Robey & Barcikowski, 1992).

The relative performance of the analysis strategies was evaluated by a comparison of the Type I error control and statistical power of the tests of regression coefficients. Estimates of Type I error control and statistical power were obtained by conducting the hypothesis tests associated with each parameter of each analytical model. For the OLS and LR models, the simultaneous test that all regression parameters is equal to zero was also conducted. The proportion of rejections of hypothesis tests for conditions in which

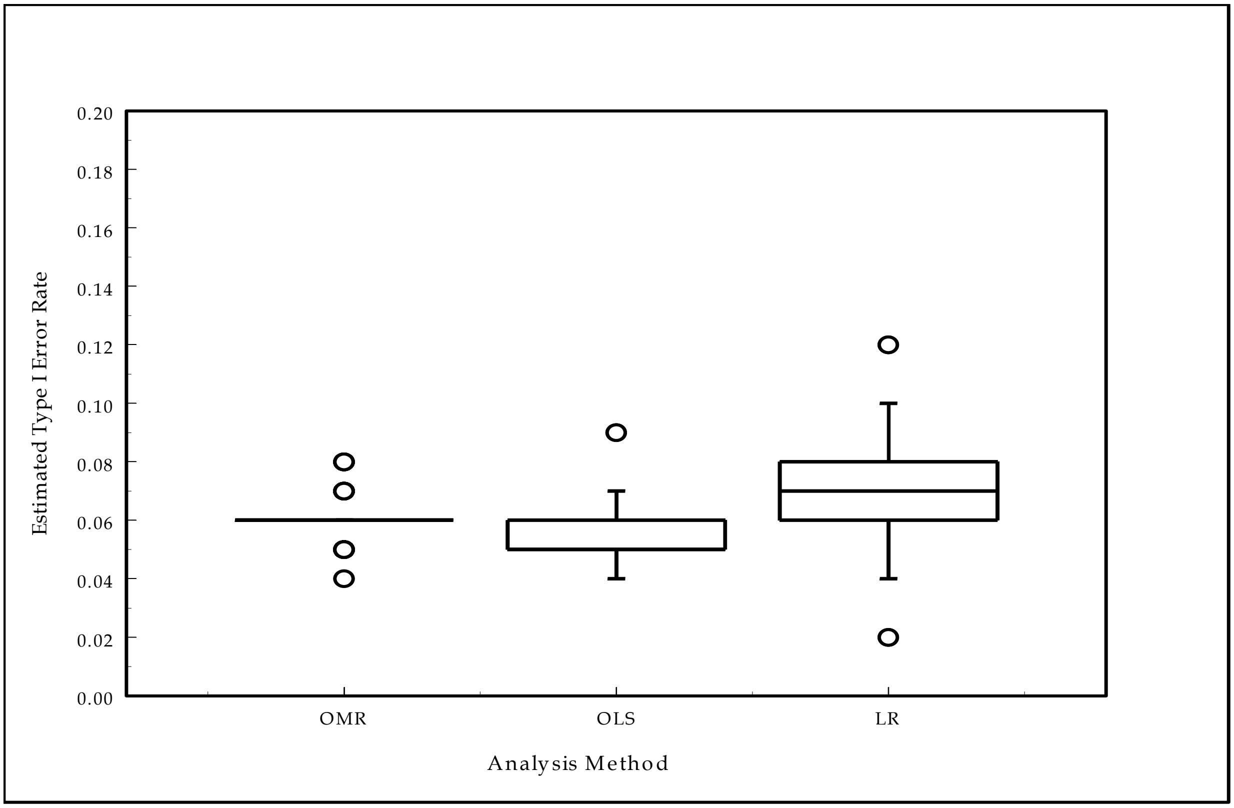

FIGURE 1: Distribution of Type I Error Rate Estimates for Tests of Individual Regression Coefficients

the population parameter is zero provided an estimate of Type I error control. Similarly, the proportion of rejections of such tests for conditions in which the parameter is not zero provided an estimate of statistical power.

Results

Type I Error Control

Box-and-whisker plots in Figure 1 depict the distribution of Type I error rate estimates for tests of individual regression weights across the conditions examined in the research. The figure clearly portrays a higher typical Type I error rate for the LR analysis method, relative to the OMR and OLS methods. The tails of the plot for the LR method range from .02 to .12, with half ranging from .06 to .08. The OMR method exhibits the most consistent Type I estimates, with the majority of estimates at .06. The OLS method is similar to the OMR with tails ranging from .04 to .07, with half ranging from .05 to .06.

Estimates of Type I error rates for the tests of individual regression weights in samples from the unimodal symmetric population are presented in Table 1. Overall the 5-point scale and 7-point scale exhibited few differences across the three analysis methods for all conditions. With two regressors and small sample size, the Type I error rate for the LR analysis method was lower than the OMR and OLS analysis methods. For example, with a 5-point scale, k=2, r_{12}=.30, and n=10, the estimated Type I error rate for OMR was .07, for OLS was .05, and for LR was .04. However, with more than 2 regressors and small to medium sample sizes, the Type I error rate was higher for LR when compared to OLS and OMR. For example, with a 5-point scale, k=10, r_{12}=.10, and n=50, the estimated Type I error rate for LR (.09) was higher than OMR (.05) and OLS (.05) . The OMR method had slightly higher Type I error rates with 2 regressors when compared to its performance with 5 and 10 regressors.

The analysis methods provided similar results across models with five and ten regressors, with LR consistently having higher Type I error rates than OMR or OLS. The OMR and OLS methods for small, medium, and larger sample sizes exhibited similar Type I error rates, which ranged from .05 to .07. In contrast the LR analysis method for small and medium sample sizes exhibited a higher Type I error rate,

Table 1. Type I Error Rate Estimates for Tests of Individual Regression Weights with Unimodal Symmetric Populations

5-point Scale			Analysis Method		
k	r_{12}	N	OMR	OLS	LR
2	.10	10	0.08	0.06	0.04
		20	0.07	0.05	0.06
		200	0.06	0.06	0.06
	.30	10	0.07	0.05	0.04
		20	0.06	0.05	0.06
		200	0.06	0.06	0.06
5	.10	25	0.06	0.05	0.08
		50	0.06	0.06	0.07
		500	0.06	0.06	0.06
	.30	25	0.06	0.05	0.09
		50	0.06	0.06	0.07
		500	0.06	0.06	0.06
10	.10	50	0.05	0.05	0.09
		100	0.06	0.06	0.08
		1000	0.06	0.06	0.07
	.30	50	0.05	0.05	0.10
		100	0.06	0.06	0.08
		1000	0.06	0.07	0.07
7-point Scale			Analysis Method		
k	r_{12}	N	OMR	OLS	LR
2	.10	10	0.08	0.05	0.05
		20	0.06	0.05	0.06
		200	0.06	0.05	0.06
	.30	10	0.07	0.05	0.05
		20	0.06	0.05	0.06
		200	0.06	0.06	0.05
5	.10	25	0.06	0.06	0.09
		50	0.06	0.05	0.07
		500	0.06	0.06	0.06
	.30	25	0.06	0.05	0.09
		50	0.06	0.06	0.08
		500	0.05	0.05	0.05
10	.10	50	0.06	0.05	0.05
		100	0.06	0.06	0.07
		1000	0.05	0.05	0.05
	.30	50	0.06	0.06	0.10
		100	0.06	0.06	0.08
		1000	0.07	0.06	0.07

which ranged from .06 to .10. For the small sample conditions, the LR method provided slightly higher Type I error rates with 10 regressors than with 5 regressors (although it remained liberal with small samples in both conditions). For example, with a 7-point scale, k=10, r_{12}=.30, and n=50, the estimated Type I error rate for LR was .10. In contrast, with k = 5, r_{12}=.30, and n = 25, the Type I error rate estimate for LR dropped to .09. The LR method evidenced better Type I error control as the sample size increased. For example, with a 7-point scale, k=5, r_{12}=.30, the LR method was .09 with the small sample, .08 with the medium sample, and .05 with the large sample.

The Type I error rate estimates for tests of individual regression weights in selected conditions across distribution shapes are presented in Table 2. The three distribution shapes (unimodal symmetric, uniform,

Table 2. Type I Error Rates Estimates for Tests of Individual Regression Weights Across Distribution Shapes.

5-point Scale			Symmetric			Uniform			Skewed		
k	r_{12}	N	OMR	OLS	LR	OMR	OLS	LR	OMR	OLS	LR
2	.10	10	0.08	0.06	0.04	0.07	0.04	0.05	0.08	0.05	0.02
		20	0.07	0.05	0.06	0.07	0.05	0.07	0.07	0.05	0.05
	.30	10	0.07	0.05	0.04	0.07	0.06	0.05	0.08	0.05	0.02
		20	0.06	0.05	0.06	0.06	0.05	0.06	0.07	0.05	0.05
5	.10	25	0.06	0.05	0.08	0.06	0.06	0.09	0.06	0.05	0.08
		50	0.06	0.06	0.07	0.06	0.05	0.07	0.06	0.05	0.07
	.30	25	0.06	0.05	0.09	0.06	0.05	0.09	0.06	0.06	0.07
		50	0.06	0.06	0.07	0.07	0.06	0.08	0.06	0.05	0.06
10	.10	50	0.05	0.05	0.09	0.05	0.05	0.09	0.06	0.05	0.09
		100	0.06	0.06	0.08	0.05	0.05	0.07	0.05	0.05	0.07
	.30	50	0.05	0.05	0.10	0.05	0.05	0.09	0.04	0.09	0.12
		100	0.06	0.06	0.08	0.05	0.05	0.07	0.05	0.05	0.07

7-point Scale			Symmetric			Uniform			Skewed		
k	r_{12}	N	OMR	OLS	LR	OMR	OLS	LR	OMR	OLS	LR
2	.10	10	0.08	0.05	0.05	0.08	0.05	0.07	0.07	0.04	0.04
		20	0.06	0.05	0.06	0.06	0.05	0.07	0.06	0.05	0.06
	.30	10	0.07	0.05	0.05	0.08	0.05	0.07	0.07	0.06	0.04
		20	0.06	0.05	0.06	0.06	0.05	0.06	0.07	0.05	0.06
5	.10	25	0.06	0.06	0.09	0.05	0.05	0.09	0.06	0.05	0.08
		50	0.06	0.05	0.07	0.06	0.05	0.07	0.06	0.06	0.07
	.30	25	0.06	0.05	0.09	0.05	0.05	0.09	0.06	0.06	0.08
		50	0.06	0.06	0.08	0.06	0.05	0.07	0.06	0.05	0.07
10	.10	50	0.06	0.05	0.05	0.05	0.05	0.09	0.06	0.06	0.10
		100	0.06	0.06	0.07	0.06	0.06	0.07	0.06	0.06	0.08
	.30	50	0.06	0.06	0.10	0.05	0.05	0.10	0.05	0.09	0.12
		100	0.06	0.06	0.08	0.06	0.06	0.08	0.06	0.06	0.07

and skewed) yielded similar Type I error rates for all three analysis methods. For example, with the 7-point scale, k=5, r_{12}=.10, and n=25, the estimated Type I error rate ranged from .05 to .06 across the distribution shapes for the OMR and OLS analysis method. Similarly, for the LR analysis method the estimated Type I error rate ranged from .08 to .09 across the distribution shapes.

Figure 2 contains box-and-whisker plots depicting the distribution of Type I error rate estimates for simultaneous tests of regression coefficients. The LR method portrays a wide range of estimates, .00 to .18, with half ranging from .00 to .09. The OLS method portrays a much smaller range of estimates, .04 to .07, with half ranging from .05 to .06. Thus the OLS method exhibits a much smaller type I error rate for simultaneous tests of regression coefficients across all research conditions.

Table 3 presents the estimates of Type I error rates for the simultaneous test of all regression weights when samples were drawn from the unimodal symmetric populations. As evident in this table, the LR method evidenced substantially greater variability in Type I error control for this test than did the OLS method. Specifically, the LR method was very conservative in conditions with both k = 2 and k = 5 and was liberal with k = 10, with Type I error rates reaching as high as .14 (with the 7-point scale, k = 10, r_{12}=.3, and n=50). The OLS method maintained Type I error rates near the nominal level across these conditions, with larger estimates in the larger sample size and 10k conditions. The Type I error rates for these tests across distribution shapes (Table 4) indicated a slight increase in Type I error rates for the OLS method when samples were drawn from the 7 point, 10 regressors, symmetric and skewed distributions, relative to the rates obtained from the uniform distribution. However, these higher rates remained very close to the nominal alpha level and did not exceed .09.

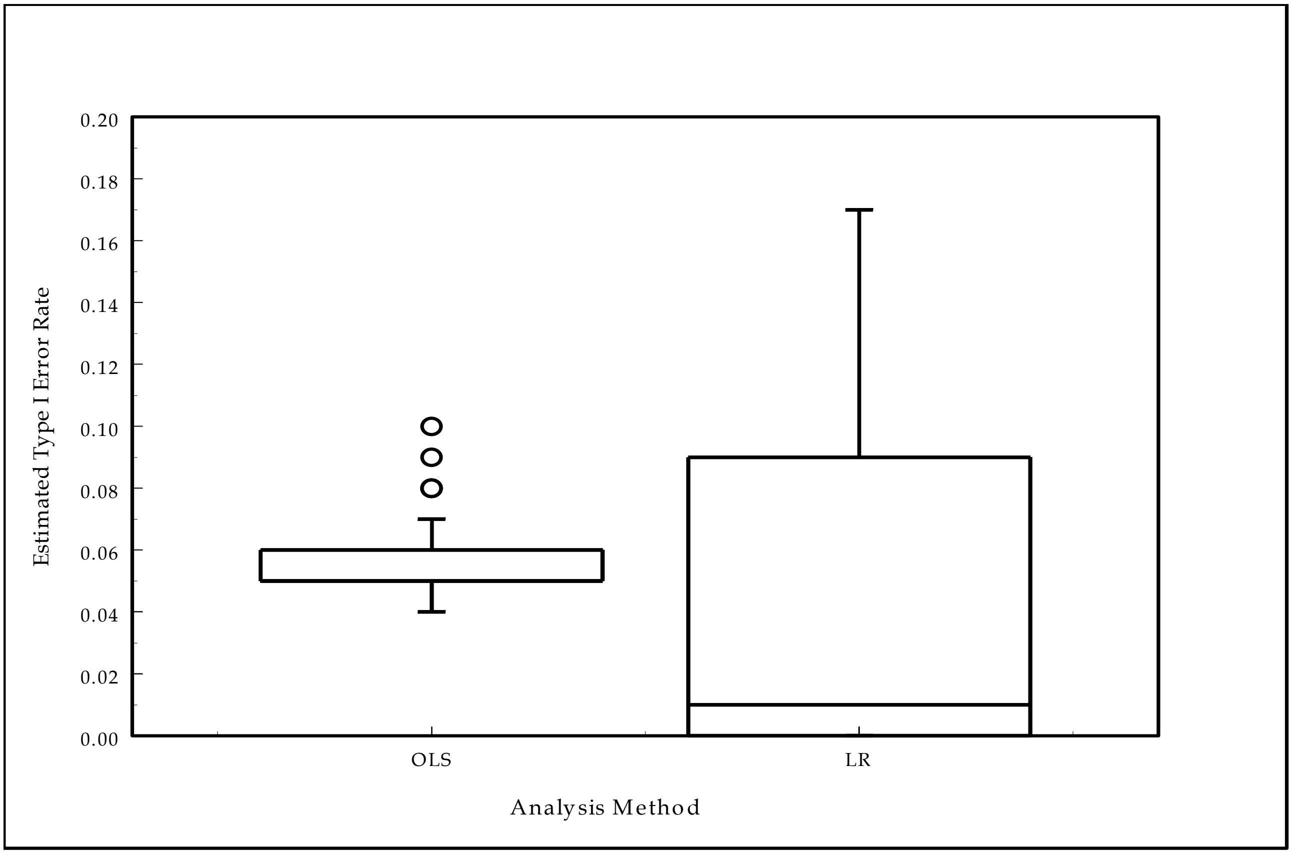

FIGURE 2: Distribution of Type I Error Rate Estimates for Simultaneous Test of Regression Coefficients

Statistical Power

The distribution of estimates of statistical power for the tests of regression weights across the conditions examined in this research are presented as box-and-whisker plots in Figure 3. As is evident in this figure, most of the conditions examined yielded relatively low power and the power differences across analysis methods was not great. However, the LR analysis method provided slight power advantages relative to OMR and OLS analyses (a result that should be expected because of its tendency to be liberal in Type I error control).

Power estimates for the tests of individual regression weights with samples drawn from the three distribution shapes are presented in Tables 5 and 6. Few differences were evident between the analyses of variables with 5 points and those with 7 points. With small effect sizes (f^2 = .02) the power estimates for all analysis methods ranged from .04 to .12. With large effect sizes (f^2 = .35), modest but greater power differences were evident across the analysis methods. With a small number of regressors (k = 2), the OMR method evidenced power advantages over the OLS and LR methods with small samples. With a 5-point scale, f^2 = .35, k = 2, r_{12} = .10, and n = 10, for example, the statistical power of the OMR method was estimated to be .19, while the power estimates of the OLS and LR methods were .15 and .09, respectively (Table 5). These slight power advantages of the OMR method disappeared with larger samples and larger numbers of regressors (conditions in which the LR approach evidenced modest power advantages). The differences in statistical power between tests conducted on samples from unimodal symmetric, uniform, and skewed distributions are negligible. For example, with the 7-point response scale, f^2 = .35, k = 5, r_{12} = .10, and n = 50, the power estimates for the OMR method ranged from .21 for the skewed population to .28 for the uniform population (Table 6). Similarly, the power estimates for OLS ranged only from .28 to .31, and those of the LR method ranged from .32 to .35.

The distribution of power estimates for simultaneous tests of regression coefficients across all research conditions is presented with box-and-whisker plots in Figure 4. The OLS method evidenced a slight power advantage over the LR method. Power estimates for the simultaneous test of all regression

Table 3. Estimates of Type I Error Rate for Simultaneous Tests of All Regression Weights with Unimodal Symmetric Populations

			5-Point Scale Analysis Method		7-Point Scale Analysis Method	
k	r_{12}	N	OLS	LR	OLS	LR
2	.10	10	0.05	0.01	0.05	0.01
		20	0.05	0.00	0.04	0.00
		200	0.06	0.00	0.06	0.00
	.30	10	0.05	0.00	0.05	0.01
		20	0.05	0.00	0.05	0.00
		200	0.05	0.00	0.06	0.00
5	.10	25	0.05	0.01	0.06	0.01
		50	0.06	0.01	0.06	0.01
		500	0.08	0.00	0.07	0.00
	.30	25	0.06	0.01	0.05	0.02
		50	0.05	0.01	0.08	0.01
		500	0.08	0.01	0.05	0.00
10	.10	50	0.06	0.11	0.07	0.13
		100	0.06	0.10	0.06	0.09
		1000	0.08	0.08	0.08	0.08
	.30	50	0.06	0.14	0.07	0.14
		100	0.07	0.10	0.07	0.10
		1000	0.10	0.10	0.08	0.09

weights with samples drawn from unimodal symmetric, uniform, and skewed populations are presented in Tables 7 and 8. With a small effect size (.02) and k = 2 or k = 5, the LR test provided rejection rates less than the nominal alpha level across most conditions. For the OLS test, similarly low power levels were observed except for large samples (and power estimates were still less than .50 for these conditions). As expected, with larger effect sizes the power estimates for both methods increased. The OLS method evidenced substantially greater power than the LR method for k = 2 and k = 5, while the LR method evidenced smaller power advantages for k = 10. A comparison of power estimates across distribution shapes suggests that population shape is not a major influence on the power of either the OLS or the LR method.

Conclusions

Surprisingly few and relatively small differences were evident among the OMR, OLS and LR methods in terms of their Type I error control and statistical power in tests of their respective weights. The OMR approach, while accurately representing the ordinal nature of the discrete response scales and discrete regressor variables, provided neither superior Type I error control nor superior statistical power. In contrast, the LR analysis evidenced Type I error control problems with small sample sizes or large numbers of regressors, conditions in which the Wald tests of the logistic regression weights became liberal. Although this analysis method presents modest power advantages relative to OMR and OLS, such power comes at a cost of relatively tenuous Type I error control. Finally, the surprisingly good performance of the OLS approach suggests that researchers who approach the analysis of discrete ordinal data (such as individual Likert items) with OLS tools should feel no guilt in such a tactic. The Type I error control in tests conducted of the OLS regression weights was as good as that obtained with tests of OMR weights and was superior to tests obtained in the LR context. Further, the statistical power evidenced with OLS was comparable or superior to that of OMR.

Of course, testing hypotheses about weights obtained in these models are only a small part of the inferential machinery applied to discrete ordinal data. Additional research is needed to focus on the relative bias in sample estimates of weights for OMR, OLS, and LR models and on the accuracy of confidence bands constructed around the sample estimated weights. Furthermore, research conducted on models that present a mixture of discrete ordinal variables and continuous variables is needed in order to explore the relative performance of these models in such a context. Finally, the substantive nature of the

Table 4. Type I Error Rate Estimates for Simultaneous Tests of All Regression Weights Across Distribution Shapes.

5-point Scale			Symmetric		Uniform		Skewed	
k	r_{12}	N	OLS	LR	OLS	LR	OLS	LR
2	.10	10	0.05	0.01	0.05	0.00	0.05	0.00
		20	0.05	0.00	0.06	0.00	0.05	0.00
	.30	10	0.05	0.00	0.06	0.01	0.05	0.01
		20	0.05	0.00	0.05	0.00	0.05	0.00
5	.10	25	0.05	0.01	0.06	0.02	0.05	0.01
		50	0.06	0.01	0.06	0.00	0.05	0.01
	.30	25	0.06	0.01	0.06	0.01	0.05	0.01
		50	0.05	0.01	0.06	0.01	0.05	0.01
10	.10	50	0.06	0.11	0.05	0.12	0.05	0.12
		100	0.06	0.10	0.05	0.09	0.06	0.09
	.30	50	0.06	0.14	0.04	0.13	0.04	0.14
		100	0.07	0.10	0.05	0.09	0.05	0.12

7-point Scale			Symmetric		Uniform		Skewed	
k	r_{12}	N	OLS	LR	OLS	LR	OLS	LR
2	.10	10	0.05	0.01	0.05	0.00	0.05	0.00
		20	0.04	0.00	0.05	0.00	0.05	0.00
	.30	10	0.05	0.01	0.05	0.01	0.06	0.01
		20	0.05	0.00	0.05	0.00	0.05	0.00
5	.10	25	0.06	0.01	0.04	0.01	0.04	0.01
		50	0.06	0.01	0.05	0.00	0.06	0.01
	.30	25	.0.05	0.02	0.05	0.01	0.06	0.02
		50	0.08	0.01	0.05	0.00	0.06	0.01
10	.10	50	0.07	0.13	0.05	0.12	0.07	0.15
		100	0.06	0.09	0.06	0.09	0.09	0.12
	.30	50	0.07	0.14	0.05	0.15	0.05	0.17
		100	0.07	0.10	0.05	0.09	0.06	0.13

inferences suggested by these models requires attention from the perspective of psychometrics. That is, the OMR model provides a vehicle for inferences about the ordinal position of observations on the criterion variable, while the LR model estimates the probability of an observation being "in category j or lower" on the criterion variable. In contrast, the OLS model provides a prediction of the criterion value as though it was a continuous variable (i.e., seeking to minimize the sum of squared errors of prediction). In addition (although not analyzed in this research study), methods for determining both the strength and direction of the response have been proposed (Jones & Sobel (2000); Brody & Dietz (1997) . These types of inferences have obvious substantive differences, and the validity of such inferences extends beyond the estimation of Type I error rates and statistical power.

In conclusion, this study has sought to bring increased awareness and clarity to three analysis options for multiple regression with discrete ordinal variables. The social sciences often focus on variables which are measured with ordinal scales (most commonly Likert scales). Unfortunately, the appropriate application of multiple regression with discrete ordinal data has been insufficiently addressed in the literature. There are many reasons for this paucity of treatment of the issue: general debates continue about the treatment of ordinal data with statistics, the breadth of analysis options are spread among numerous sources, and the consequences of analysis choice in terms of Type I error control and statistical power has not been thoroughly investigated. The current study presents three analysis options within the literature and clarifies their differences with the hopes of both increasing the appropriate analysis of discrete ordinal variables and stimulating additional methodological research in this area.

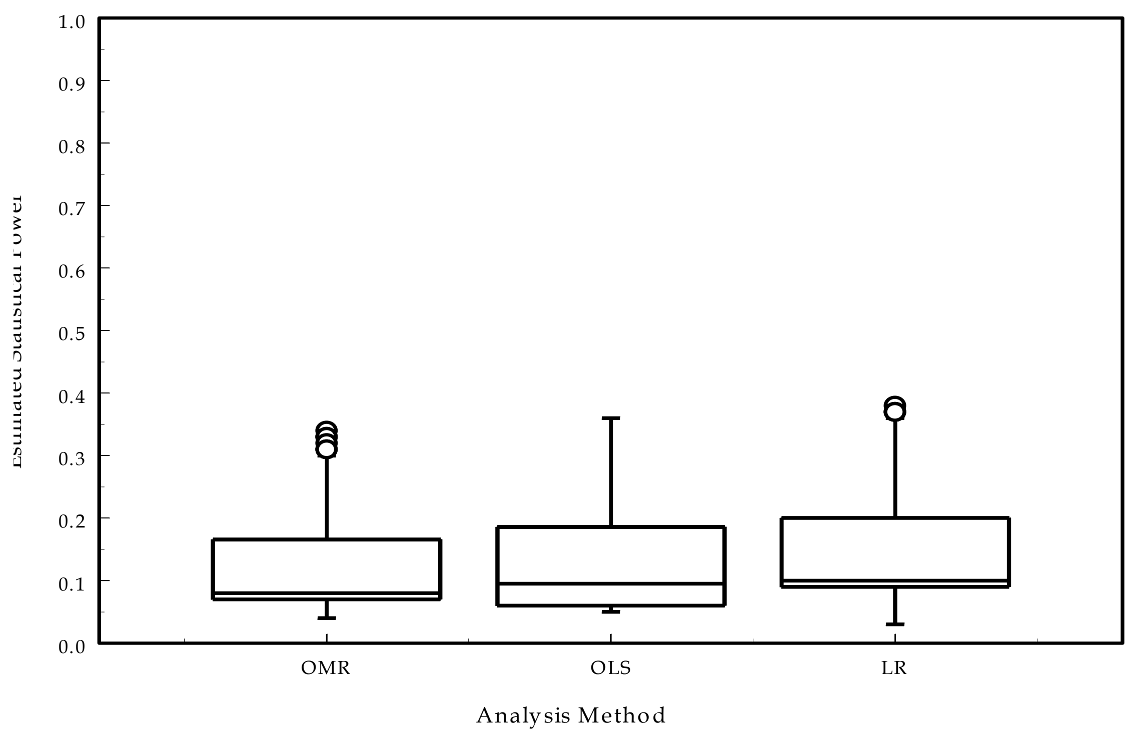

FIGURE 3: Distribution of Statistical Power Estimates for Tests of Individual Regression Coefficients

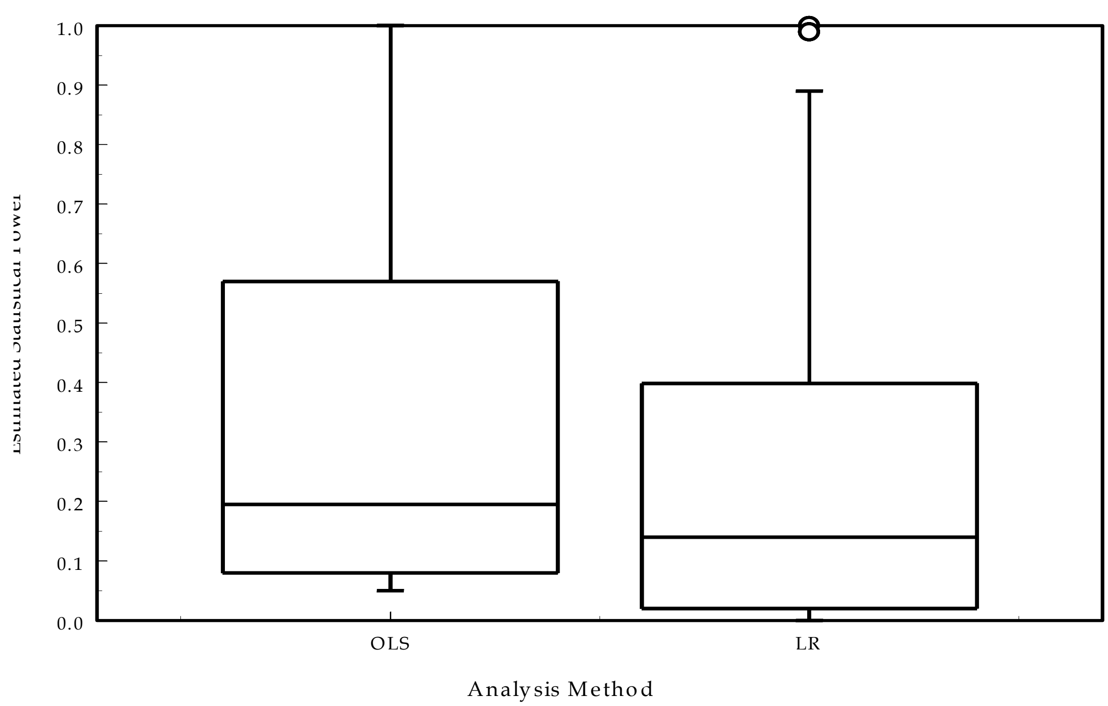

FIGURE 4: Distribution of Statistical Power Estimates for Simultaneous Test of Regression Coefficients

Table 5. Estimates of Statistical Power for Tests of Individual Regression Weights Across Distribution Shapes.

5 Point Likert				Symmetric			Uniform			Skewed		
Effect	k	r_{12}	N	OMR	OLS	LR	OMR	OLS	LR	OMR	OLS	LR
.02	2	.10	10	0.09	0.06	0.04	0.09	0.05	0.04	0.08	0.06	0.03
			20	0.09	0.07	0.09	0.08	0.06	0.08	0.08	0.07	0.07
		.30	10	0.09	0.06	0.04	0.09	0.05	0.03	0.08	0.07	0.03
			20	0.08	0.07	0.08	0.10	0.08	0.10	0.07	0.08	0.08
	5	.10	25	0.07	0.06	0.09	0.07	0.06	0.10	0.06	0.06	0.09
			50	0.08	0.07	0.09	0.08	0.07	0.09	0.07	0.07	0.08
		.30	25	0.06	0.05	0.09	0.07	0.05	0.09	0.05	0.11	0.11
			50	0.11	0.10	0.12	0.06	0.06	0.07	0.05	0.08	0.08
	10	.10	50	0.07	0.06	0.10	0.06	0.06	0.10	0.06	0.06	0.10
			100	0.07	0.07	0.09	0.07	0.06	0.08	0.06	0.07	0.09
		.30	50	0.06	0.05	0.09	0.05	0.05	0.09	0.04	0.10	0.15
			100	0.07	0.06	0.08	0.06	0.06	0.08	0.05	0.07	0.09
.35	2	.10	10	0.19	0.15	0.09	0.18	0.16	0.12	0.18	0.19	0.08
			20	0.34	0.34	0.37	0.31	0.32	0.35	0.28	0.35	0.36
		.30	10	0.19	0.14	0.07	0.17	0.14	0.10	0.14	0.18	0.07
			20	0.30	0.27	0.30	0.33	0.31	0.34	0.25	0.31	0.30
	5	.10	25	0.15	0.16	0.22	0.13	0.14	0.20	0.13	0.19	0.23
			50	0.27	0.30	0.34	0.24	0.26	0.29	0.22	0.33	0.36
		.30	25	0.11	0.11	0.16	0.10	0.10	0.15	0.07	0.19	0.17
			50	0.20	0.20	0.23	0.18	0.18	0.21	0.11	0.23	0.22
	10	.10	50	0.11	0.14	0.19	0.11	0.13	0.19	0.09	0.15	0.19
			100	0.21	0.24	0.28	0.21	0.24	0.28	0.16	0.25	0.27
		.30	50	0.07	0.08	0.14	0.05	0.08	0.14	0.04	0.19	0.19
			100	0.11	0.12	0.15	0.09	0.12	0.17	0.05	0.18	0.17

References

Agresti, A. & Finlay, B. (1997). *Statistical methods for the social sciences* (3rd Ed.). Upper Saddle River, NJ: Prentice-Hall.

Borgatta, E. F. & Bohrnstedt, G. W. (1980). Level of measurement: Once over again. *Sociological Methods & Research*, *9*(2), 147-160.

Bradley, D. R. & Fleisher, C. L. (1994). Generating multivariate data from nonnormal distributions: Mihal and Barrett revisited. *Behavior Research Methods, Instruments, and Computers*, *26*, 156-166.

Brody, C. J. & Dietz, J. (1997). On the dimensionality of two-question format likert attitude scales. *Social Science Research*, *26*, 197-204.

Cliff, N. (1994). Predicting ordinal relations. *British Journal of Mathematical and Statistical Psychology*, *47*, 127-150.

Cliff, N. (1996). *Ordinal Methods for Behavioral Data Analysis*. Mahwah, NJ: Lawrence Erlbaum Associates.

Ferron, J., Yi, Q. & Kromrey, J.D. (1997). NNCORR: A SAS/IML program for generating nonnormal correlated data. *Applied Psychological Measurement*, *21*, 64.

Table 6. Estimates of Statistical Power for Tests of Individual Regression Weights Across Distribution Shapes.

7 point Likert				Symmetric			Uniform			Skewed		
Effect	k	r_{12}	N	OMR	OLS	LR	OMR	OLS	LR	OMR	OLS	LR
.02	2	.10	10	0.08	0.06	0.06	0.08	0.05	0.07	0.08	0.06	0.04
			20	0.08	0.07	0.09	0.08	0.07	0.09	0.07	0.07	0.08
		.30	10	0.08	0.05	0.05	0.08	0.06	0.07	0.09	0.07	0.04
			20	0.08	0.07	0.08	0.08	0.06	0.08	0.09	0.08	0.08
	5	.10	25	0.07	0.06	0.10	0.07	0.06	0.10	0.07	0.08	0.11
			50	0.07	0.07	0.09	0.07	0.07	0.09	0.07	0.08	0.09
		.30	25	0.06	0.05	0.09	0.06	0.05	0.09	0.05	0.09	0.11
			50	0.07	0.07	0.09	0.06	0.06	0.08	0.06	0.07	0.10
	10	.10	50	0.06	0.06	0.11	0.06	0.06	0.10	0.06	0.06	0.10
			100	0.08	0.07	0.09	0.07	0.07	0.09	0.07	0.07	0.09
		.30	50	0.06	0.05	0.09	0.05	0.06	0.10	0.04	0.14	0.18
			100	0.06	0.06	0.08	0.06	0.06	0.08	0.04	0.09	0.09
.35	2	.10	10	0.18	0.15	0.12	0.18	0.16	0.17	0.16	0.20	0.10
			20	0.32	0.33	0.37	0.33	0.34	0.38	0.26	0.36	0.37
		.30	10	0.18	0.13	0.09	0.17	0.14	0.14	0.12	0.15	0.10
			20	0.31	0.28	0.31	0.29	0.28	0.32	0.25	0.32	0.32
	5	.10	25	0.14	0.14	0.20	0.14	0.16	0.23	0.12	0.18	0.22
			50	0.26	0.28	0.32	0.28	0.31	0.35	0.21	0.31	0.34
		.30	25	0.12	0.11	0.16	0.11	0.11	0.17	0.07	0.19	0.18
			50	0.20	0.18	0.22	0.19	0.19	0.22	0.13	0.24	0.24
	10	.10	50	0.11	0.13	0.19	0.11	0.13	0.19	0.10	0.16	0.21
			100	0.21	0.23	0.28	0.22	0.25	0.28	0.18	0.26	0.28
		.30	50	0.07	0.08	0.14	0.07	0.09	0.16	0.04	0.19	0.20
			100	0.12	0.13	0.16	0.12	0.13	0.18	0.05	0.19	0.19

Gaito, J. (1980). Measurement scales and statistics: Resurgence of an old misconception. *Psychological Bulletin*, *87*(*3*), 564-567.

Harwell, M. R. & Gatti, G. G. (2001). Rescaling ordinal data to interval data in educational research. *Review of Educational Research*, *71*, 105–131.

Jones, B. S. & Sobel, M. E. (2000). Modeling direction and intensity in semantically balanced ordinal scales: An assessment of congressional incumbent approval. *American Journal of Political Science*, *44*(*1*), 174-185.

Long, J. D. (1999). A confidence interval for ordinal multiple regression weights. *Psychological Methods*, *4*(*3*), 315-330.

Stevens, S. S. (1951). Mathematics, measurement and psychophysics. In S. S. Stevens (Ed.), *Handbook of experimental psychology* 1-49. New York: Wiley.

Robey, R. R. & Barcikowski, R. S. (1992). Type I error and the number of iterations in Monte Carlo studies of robustness. *British Journal of Mathematical and Statistical Psychology*, *45*, 283-288.

Table 7. Estimates of Statistical Power for Simultaneous Tests of All Regression Weights Across Distribution Shapes.

5 Point Likert				Symmetric		Uniform		Skewed	
Effect	k	r_{12}	N	OLS	LR	OLS	LR	OLS	LR
.02	2	.10	10	0.07	0.01	0.05	0.01	0.07	0.01
			20	0.08	0.00	0.08	0.00	0.07	0.00
		.30	10	0.06	0.01	0.05	0.01	0.08	0.01
			20	0.07	0.00	0.09	0.00	0.09	0.00
	5	.10	25	0.06	0.01	0.08	0.03	0.08	0.02
			50	0.11	0.01	0.13	0.02	0.12	0.01
		.30	25	0.06	0.02	0.06	0.02	0.10	0.02
			50	0.35	0.18	0.08	0.01	0.13	0.02
	10	.10	50	0.10	0.18	0.08	0.16	0.11	0.17
			100	0.14	0.18	0.13	0.17	0.15	0.17
		.30	50	0.09	0.20	0.06	0.19	0.17	0.19
			100	0.15	0.22	0.10	0.16	0.17	0.23
.35	2	.10	10	0.21	0.02	0.22	0.03	0.28	0.02
			20	0.53	0.04	0.50	0.03	0.53	0.04
		.30	10	0.22	0.03	0.22	0.02	0.26	0.02
			20	0.52	0.04	0.55	0.04	0.54	0.04
	5	.10	25	0.44	0.20	0.42	0.18	0.54	0.24
			50	0.85	0.56	0.80	0.49	0.87	0.58
		.30	25	0.51	0.27	0.38	0.18	0.52	0.22
			50	0.90	0.66	0.84	0.52	0.86	0.52
	10	.10	50	0.74	0.87	0.72	0.84	0.73	0.82
			100	0.99	1.00	0.99	0.99	0.98	0.99
		.30	50	0.69	0.84	0.71	0.89	0.75	0.81
			100	0.98	0.99	0.99	1.00	0.99	0.99

Table 8. Estimates of Statistical Power for Simultaneous Tests of All Regression Weights Across Distribution Shapes.

7 Point Likert				Symmetric		Uniform		Skewed	
Effect	k	r_{12}	N	OLS	LR	OLS	LR	OLS	LR
.02	2	.10	10	0.06	0.00	0.06	0.01	0.07	0.01
			20	0.08	0.00	0.09	0.00	0.09	0.00
		.30	10	0.06	0.01	0.06	0.00	0.08	0.00
			20	0.08	0.00	0.08	0.00	0.12	0.00
	5	.10	25	0.07	0.02	0.07	0.02	0.14	0.03
			50	0.11	0.02	0.10	0.02	0.13	0.02
		.30	25	0.06	0.02	0.08	0.02	0.10	0.03
			50	0.11	0.02	0.09	0.01	0.13	0.02
	10	.10	50	0.08	0.17	0.08	0.19	0.11	0.18
			100	0.13	0.20	0.14	0.20	0.15	0.19
		.30	50	0.06	0.14	0.06	0.18	0.08	0.18
			100	0.10	0.14	0.10	0.16	0.19	0.22
.35	2	.10	10	0.20	0.02	0.22	0.02	0.27	0.02
			20	0.51	0.03	0.52	0.03	0.54	0.04
		.30	10	0.20	0.02	0.23	0.02	0.23	0.03
			20	0.53	0.05	0.51	0.04	0.56	0.05
	5	.10	25	0.40	0.18	0.46	0.21	0.49	0.22
			50	0.81	0.52	0.86	0.58	0.84	0.55
		.30	25	0.45	0.23	0.44	0.21	0.56	0.25
			50	0.85	0.58	0.86	0.55	0.88	0.55
	10	.10	50	0.68	0.82	0.70	0.82	0.76	0.84
			100	0.98	0.99	0.99	1.00	0.98	0.99
		.30	50	0.74	0.88	0.76	0.89	0.82	0.87
			100	0.99	1.00	1.00	1.00	0.99	1.00

Send correspondence to: Jeffrey D. Kromrey
Department of Educational Measurement and Research
University of South Florida
4202 E. Fowler Ave., EDU162, Tampa, FL 33620
Email: Kromrey@tempest.coedu.usf.edu