QUANTITATIVE METHODOLOGY SERIES
Methodology for Business and Management

George A. Marcoulides, Series Editor

The purpose of this series is to present methodological techniques to investigators and students from all functional areas of business, although individuals from other disciplines will also find the series useful. Each volume in the series will focus on a specific method (e.g., Data Envelopment Analysis, Factor Analysis, Multilevel Analysis, Structural Equation Modeling). The goal is to provide an understanding and working knowledge of each method with a minimum of mathematical derivations.

Proposals are invited from all interested authors. Each proposal should consist of the following: (i) a brief description of the volume's focus and intended market, (ii) a table of contents with an outline of each chapter, and (iii) a curriculum vita. Materials may be sent to Dr. George A. Marcoulides, Department of Management Science, California State University, Fullerton, CA 92634.

Marcoulides • Modern Methods for Business Research

Duncan/Duncan/Strycker/Li/Alpert • An Introduction to Latent Variable Growth Curve Modeling: Concepts, Issues, and Applications

Heck/Thomas • An Introduction to Multilevel Modeling Techniques

Marcoulides/Moustaki • Latent Variable and Latent Structure Models

Hox • Multilevel Analysis: Techniques and Applications

Multilevel Analysis

Techniques and Applications

Joop Hox
Utrecht University, the Netherlands

2002
LAWRENCE ERLBAUM ASSOCIATES, PUBLISHERS
Mahwah, New Jersey London

the most experienced teacher is (25-2=) 23×0.11=2.53 points on the popularity measure. We can use the standard errors of the regression coefficients reported in Table 2.1 to construct a 95% confidence interval. For the regression coefficient of pupil gender, the 95% confidence interval runs from 0.72 to 0.96, and the 95% confidence interval for the regression coefficient of teacher experience runs from 0.09 to 0.13.

The model with the explanatory variables includes a variance component for the regression coefficient of pupil gender, symbolized by σ^2_{u1} in Table 2.1. The variance of the regression coefficients for pupil gender across classes is estimated as 0.27, with a standard error of 0.05. The covariance between the regression coefficient for pupil gender and the intercept is very small and obviously not significant.

The significant and quite large variance of the regression slopes for pupil gender implies that we should not interpret the estimated value of 0.84 without considering this variation. In an ordinary regression model, without multilevel structure, the value of 0.84 means that girls are expected to differ from boys by 0.84 points, for all pupils in all classes. In our multilevel model, the regression coefficient for pupil gender varies across the classes, and the value of 0.84 is just the expected value across all classes. In multilevel regression analysis, the varying regression coefficients are assumed to follow a normal distribution. The variance of this distribution is in our example estimated as 0.27. Interpretation of this variation is easier when we consider the standard deviation, which is the square root of the variance or 0.52 in our example data. A useful characteristic of the standard deviation is that with normally distributed observations about 67% of the observations lie between one standard deviation below and above the mean, and about 95% of the observations lie between two standard deviations below and above the mean. If we apply this to the regression coefficients for pupil gender, we conclude that about 67% of the regression coefficients are expected to lie between (0.84-0.52=) 0.32 and (0.84+0.52=) 1.36, and about 95% are expected to lie between (0.84-1.04=) –0.20 and (0.84+1.04=) 1.88. Using the more precise value of $Z_{.975}$=1.96 we calculate the limits of the 95% interval as –0.18 and 1.86. We can also use the standard normal distribution to estimate the percentage of regression coefficients that are negative. As it turns out, even if the mean regression coefficient for pupil gender is 0.84, about 5% of the classes are expected to have a regression coefficient that is actually negative. Note that the 95% interval computed here is totally different from the 95% confidence interval for the regression coefficient of pupil gender, which runs from 0.72 to 0.96. The 95% confidence interval applies to γ_{10}, the mean value of the regression coefficients across the classes. The 95% interval calculated here is the 95% *predictive interval*, which expresses that 95% of the regression coefficients of the variable 'pupil gender' in the classes are predicted to lie between –0.20 and 1.88.

Given the large and significant variance of the regression coefficient of pupil gender across the classes it is attractive to attempt to predict its variation using class level variables. We have one class level variable: teacher experience. The individual level regression equation for this example, using variable labels instead of symbols, is given by equation (2.2), which is repeated below:

$$popularity_{ij} = \beta_{0j} + \beta_{1j}\, gender_{ij} + e_{ij} \qquad \text{(2.2, repeated)}$$

The regression equations predicting β_{0j}, the intercept in class j, and β_{1j}, the regression slope of pupil gender in class j, are given by equation (2.3) and (2.4), which are rewritten below using variable labels

$$\beta_{0j} = \gamma_{00} + \gamma_{01}\, t.exp_j + u_{0j} \qquad (2.10)$$

$$\beta_{1j} = \gamma_{10} + \gamma_{11}\, t.exp_j + u_{1j} \qquad (2.11)$$

By substituting (2.10) and (2.11) into (2.2) we get

$$\begin{aligned} popularity_{ij} = {} & \gamma_{00} + \gamma_{10}\, gender_{ij} + \gamma_{01}\, t.exp_j + \gamma_{11}\, gender_{ij} \times t.exp_j \\ & + u_{1j}\, gender_{ij} + u_{0j} + e_{ij} \end{aligned} \qquad (2.12)$$

The algebraic manipulations of the equations above make clear that to explain the variance of the regression coefficients β_{1j}, we need to introduce an interaction term in the model. This interaction, between the variables pupil gender and teacher experience, is a cross-level interaction, because it involves explanatory variables from different levels. Table 2.2 presents the estimates from a model with this cross-level interaction. For comparison, the estimates for the model without this interaction are also included in Table 2.2.

The estimates for the fixed coefficients in Table 2.2 are similar for both models, except the regression slope for pupil gender, which is considerably larger in the cross-level model. The interpretation remains the same: girls are more popular than boys are. The regression coefficient for the cross-level interaction is –0.03, which is small but significant. This interaction is formed by multiplying the scores for the variables 'pupil gender' and 'teacher experience,' and the negative value means that with experienced teachers, the advantage of being a girl is smaller than expected from the direct effects only. Thus, the difference between boys and girls is smaller with more experienced teachers.

Table 2.2 Results pupil gender and teacher experience, cross-level interaction

Model:	M1: + pup. gender and t. exp.		M2: + cross-level interaction	
Fixed part				
Predictor	coefficient	standard error	coefficient	standard error
intercept	3.34	0.16	3.31	0.16
pupil gender	0.84	0.06	1.33	0.13
teacher exp.	0.11	0.01	0.11	0.01
pup. gender × teacher exp.			-.03	0.01
Random part				
σ^2_e	0.39	0.01	0.39	0.01
σ^2_{u0}	0.40	0.06	0.40	0.06
σ^2_{u1}	0.27	0.05	0.22	0.04
σ_{u01}	0.02	0.04	0.02	0.04
Deviance	4261.2		4245.9	

Comparison of the other results between the two models shows that the variance component for pupil gender goes down from 0.27 in the direct effects model to 0.22 in the cross-level model. Apparently, the cross-level model explains some of the variation of the slopes for pupil gender. The deviance also goes down, which indicates that the model fits better than the previous model.

The coefficients in Tables 2.1 and 2.2 are all unstandardized regression coefficients. To interpret them properly, we must take the scale of the explanatory variables into account. In multiple regression analysis, and structural equation models, for that matter, the regression coefficients are often standardized because that facilitates the interpretation when one wants to compare the effects of different variables within one sample. Only if the goal of the analysis is to compare parameter estimates from different samples to each other, should one always use unstandardized coefficients. To standardize the regression coefficients, as presented in Table 2.1 or Table 2.2, one could standardize all variables before putting them into the multilevel analysis. However, this would in general also change the estimates of the variance components. This may not be a bad thing in itself, because standardized variables are also centered on their overall mean. Centering explanatory variables has some distinct advantages, which are discussed in Chapter Four. Even so, it is also possible to derive the standardized regression coefficients from the unstandardized coefficients:

$$\text{Standardized coefficient} = \frac{\text{(unstandardized coeff.)} \times \text{(stand. dev. explanatory var.)}}{\text{stand. dev. outcome variable}} \tag{2.13}$$

In our example data, the standard deviations are: 1.23 for popularity, 0.50 for gender, and 6.55 for teacher experience. Table 2.3 presents the unstandardized and standardized coefficients for the second model in Table 2.1. It also presents the estimates that we obtain if we first standardize all variables, and then carry out the analysis.

Table 2.3 Comparing unstandardized and standardized estimates

Model:	Standardization after estimation			Using standardized variables	
Fixed part	unstandardized		standardized		
Predictor	coefficient	s.e.	coefficient	coefficient	s.e.
intercept	3.34	0.16	-	-	-
pupil gender	0.84	0.06	0.34	0.34	0.02
teacher exp.	0.11	0.01	0.59	0.58	0.05
Random part					
σ^2_e	0.39	0.01		0.26	0.01
σ^2_{u0}	0.40	0.06		0.32	0.05
σ^2_{u1}	0.27	0.05		0.05	0.01
σ_{u01}	0.02	0.04		0.05	70.02
Deviance	4261.2			3446.5	

Table 2.3 shows that the standardized regression coefficients are almost the same as the coefficients estimated for standardized variables. The small differences in Table 2.3 are simply rounding errors. However, if we use standardized variables in our analysis, we find very different variance components. This is not only the effect of scaling the variables differently, which becomes clear if we realize that the covariance between the slope for pupil gender and the intercept is significant for the standardized variables. This kind of difference in results is general. The fixed part of the multilevel regression model is invariant for linear transformations, just as the regression coefficients in the ordinary single-level regression model. This means that if we change the scale of our explanatory variables, the regression coefficients and the corresponding standard errors change by the same multiplication factor, and all associated p-values remain exactly the same. However, the random part of the multilevel regression model is not invariant for

linear transformations. The estimates of the variance components in the random part can and do change, sometimes dramatically. This is discussed in more detail in section 4.2 in Chapter Four. The conclusion to be drawn here is that, if we have a complicated random part, including random components for regression slopes, we should think carefully about the scale of our explanatory variables. If our only goal is to present standardized coefficients in addition to the unstandardized coefficients, applying equation (2.13) is safer than transforming our variables. On the other hand, we may estimate the unstandardized results, including the random part and the deviance, and then re-analyze the data using standardized variables, merely using this analysis as a computational trick to obtain the standardized regression coefficients without having to do hand calculations.

2.3 INSPECTING RESIDUALS

Inspection of residuals is a standard tool in multiple regression analysis to examine whether assumptions of normality and linearity are met (cf. Stevens, 1996; Tabachnick & Fidell, 1996). Multilevel regression analysis also assumes normality and linearity, and inspection of the residuals can be used for the same goal. There is one important difference from ordinary regression analysis; we have more than one residual, in fact, we have residuals for each random effect in the model. Consequently, many different residuals plots can be made.

2.3.1 Examples of Residuals Plots

The equation below represents the one-equation version of the direct effects model for our example data. This is the multilevel model without the cross-level interaction.

$$popularity_{ij} = \gamma_{00} + \gamma_{10}\, gender_{ij} + \gamma_{01}\, experience_j + u_{1j}\, gender_{ij} + u_{0j} + e_{ij}$$

In this model, we have three residual error terms: e_{ij}, u_{0j}, and u_{1j}. The e_{ij} are the residual prediction errors at the individual level, similar to the prediction errors in ordinary single-level multiple regression. A simple boxplot of these residuals will enable us to identify extreme outliers. An assumption that is usually made in multilevel regression analysis is that the variance of the residual errors is the same in all groups. This can be assessed by computing a one-way analysis of variance of the groups on the absolute values of the residuals, which is the equivalent of Levene's test for equality of variances in Analysis of Variance (Stevens, 1996). Bryk and Raudenbush (1992) describe a chi-square test that can be used for the same purpose, which is an option in the program HLM (Raudenbush, Bryk, Cheong, & Congdon, 2000).

The u_{0j} are the residual prediction errors at the group level, which can be used in ways analogous to the analysis of the individual level residuals e_{ij}. The u_{1j} are the residuals of the regression slopes across the groups. By plotting the regression slopes for the various groups, we get a visual impression of how much the regression slopes actually differ, and we may also be able to identify groups which have a regression slope that is wildly different from the others.

To test the normality assumption, we can plot standardized residuals against their normal scores. If the residuals have a normal distribution, the plot should show a straight diagonal line. Figure 2.1 is a scatterplot of the standardized level-1 residuals (denoted by 'const' in the graph) against their normal scores. The graph indicates close conformity to normality, and no extreme outliers. Similar plots can be made for the level-2 residuals.

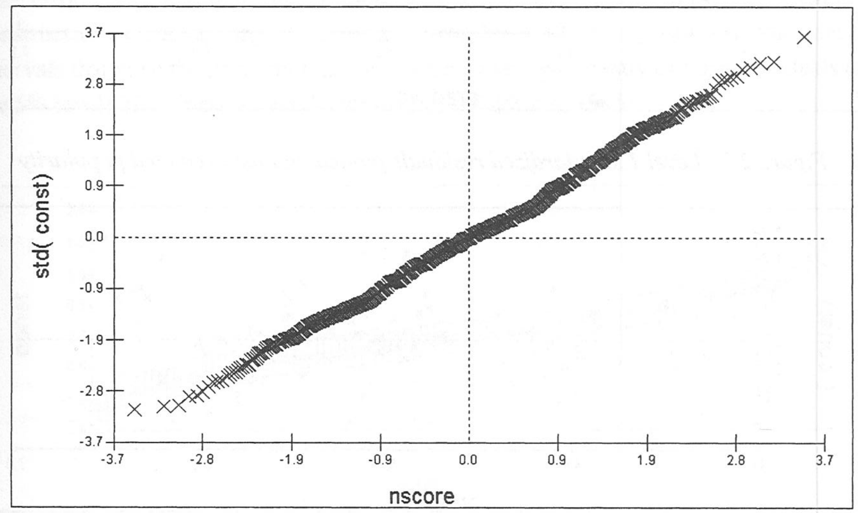

Figure 2.1. Plot of level 1 standardized residuals against normal scores

We obtain a different plot, if we plot the residuals against the predicted values of the outcome variable popularity, using the fixed part of the multilevel regression model for the prediction. Such a scatter plot of the residuals against the predicted values provides information about possible failure of normality, nonlinearity, and heteroscedasticity. If these assumptions are met, the plotted points should be evenly divided above and below their mean value of zero, with no strong structure (cf. Tabachnick & Fidell, 1996, p. 137). Figure 2.2 shows this scatter plot. For our example data, the scatter plot in Figure 2.2 does not indicate strong violations of the assumptions.