Mid-Western Educational Researcher

Volume 9
Issue 4 *Regression Analyses*

Article 3

1996

Hierarchical Modeling Techniques to Analyze Contextual Effects: What Happened to the Aptitude by Treatment Design?

Janet K. Sheehan
Northern Illinois University

Tianqi Han
Northern Illinois University

Hierarchical Modeling Techniques to Analyze Contextual Effects: What Happened to the Aptitude by Treatment Design?

Janet K. Sheehan and Tianqi Han, Northern Illinois University

Abstract

This article contrasts two analytical methods for making cross-level inferences between individual-level factors and group-level factors in school-effectiveness research, the aptitude by treatment interaction design and hierarchical linear modeling. Although the aptitude by treatment interaction method is suitable for making cross-level inferences when the intraclass correlations are low, partitioning the interaction into within-and between-contexts components is recommended to discern if the interaction is due to confounding contextual effects.

Hierarchical linear modeling is the recommended technique when intraclass correlations are high, because the parameters are assumed to be unique for each context and are modeled accordingly. Further, the shrinkage estimates of the parameters are more precise than those estimated through ordinary least squares analysis.

Contextual Effects

Educational data are often collected at the individual level, yet a cluster sampling design is used which involves first selecting groups of schools and/or classes and subsequently sampling individuals. This results in hierarchical data in which some variables are measured at the individual level and others at the group level. Since the group-level variables are nested within specific contexts such as classes, schools, or neighborhoods, it is reasonable to assume that the context itself exerts some influence on individual-level variables resulting in positive intraclass correlations. If the context is ignored in the statistical model, then there is the strong possibility of confounding effects operating at the class, school, or neighborhood level. These are known as contextual effects.

A plausible school-effects study will be considered as an example of contextual effects. In this example the purpose is to investigate the effects of instructional technology (a nominal variable reflecting multimedia and computer technology, computer technology only, and no technology) in the schools on pupil achievement. It is further expected that the effects of instructional technology will be different depending on the type of home environment the student comes from (a composite variable including such factors as the time spent talking with parents about schoolwork, reading at home, using educational computer programs, and watching educational television). The outcome is the student's composite score on a standardized achievement test. Achievement is being modeled as a function of an individual-level variable, home environment, and a group-level variable, level of instructional technology. However, there are contextual effects operating at the school level that may influence achievement other than instructional technology. The level of violence in the schools, the proximity of the school to a major university, and whether the school is public or private are a few of the other possible school-level influences on achievement. If we were able to randomly assign schools to the different instructional technologies, we could assume that over the long run, the influence of these contextual variables on achievement would be equivalent for each assigned group. However, the practice of interest is not often randomly assigned to the classes or schools in educational research, and therefore is confounded by other contextual effects.

The purpose of this paper is to contrast two analytical methods, the aptitude by treatment interaction design (ATI) and hierarchical linear modeling (HLM), for making cross-level inferences between individual-level factors and group-level factors. This article summarizes the important issues surrounding the use of each of these methods and discusses the appropriateness of each technique in school-effectiveness research.

Aptitude by Treatment Interaction

One common analysis method that has been used for models in which there is both an individual-level predictor and a group-level predictor is the ATI design. In the ATI model, the individual-level outcome (achievement) is modeled by an individual-level predictor (aptitude), a group-level treatment, and the interaction of aptitude and treatment. The interaction term is usually of most interest, because it tests whether the relationship between aptitude and achievement varies for the different treatment groups. Using the school-effects study previously mentioned, the ATI is the difference in the home environment-achievement relationship among the different levels of instructional technology.

The conventional form of the ATI parameter model that is employed in much educational research is

$$Y_{ij}=\beta_0+\beta_1X_i+\beta_2T_j+\beta_3X_iT_j+\varepsilon_{ij} \quad (1)$$

In this model each individual's outcome score is a function of aptitude (X), and treatment group (T) and the interaction of treatment and aptitude (XT). However, this model disaggregates the data to the individual level, since the context (classes or schools) are not included in the model. The major problem that disaggregation poses is that the data violate the assumption of independence of observations. Therefore positive intraclass correlations are being ignored. This is a nontrivial matter resulting in underestimated standard errors and inflated type I error in tests of significance (Bryk & Raudenbush, 1992; De Leeuw & Kreft, 1986; Draper, 1995; Cheung, Keeves, Sellin & Tsoi, 1990).

In the home environment/instructional technology example, the conventional ATI model would be appropriate if we were sure that the different schools within each type of instructional technology were of similar contexts and interacted with home environment in a consistent fashion. However, the conventional model would not be applicable if there were contextual effects. A significant interaction due to contextual effects may erroneously be attributed to other group-level variables. Conversely, the interaction of aptitude by treatment may be masked by different contexts. This occurs because in the conventional form of the ATI model we assume the β_3 to be equal for all contexts (classes or schools) within each treatment or practice. However, this assumption would often not be appropriate because of differing contextual effects in the classes or schools. In this example, the interaction of instructional technology and home environment on achievement may differ depending on another school-level variable such as the school climate (a composite variable reflecting such factors as school absenteeism, student violence, alcohol and illicit drug use, gang prevalence, weapon prevalence, and teacher abuse). Perhaps a poor school climate would mask the interaction of type of instructional technology and home environment. In schools with a good school climate, students from an impoverished home environment might benefit more from the use of instructional technology than students from an enriched home environment. However, when the school climate is poor, the use of instructional technology may not have a beneficial effect on achievement for any of the students. In this example we cannot assume that the ATI effect is consistent for different schools, because school climate masks the ATI effect in some schools.

Cronbach and Webb (1975) addressed this problem of confounding contextual effects by partitioning the ATI interaction into between-context and within-context components. The between-context component is determined by aggregating the data to the class or school level for the regression analysis. Hence the mean of Y is regressed on the mean of X, the treatment, and their interaction (2). The within-context regression is formed by first deviating the scores of the individual-level predictor from the class or school mean ($X_{ij}-\bar{X}_j$), and regressing the outcome on the deviated scores, the treatment variable, and their interaction (3).

$$\bar{Y}_j=\beta_0+\beta_1\bar{X}_j+\beta_2T_j+\beta_3\bar{X}_jT_j+\varepsilon_j \quad (2)$$

$$Y_{ij}=\beta_0+\beta_1(X_{ij}-\bar{X}_j)+\beta_2T_j+\beta_3(X_{ij}-\bar{X}_j)T_j+\varepsilon_{ij} \quad (3)$$

Cronbach and Webb demonstrated that very different conclusions can be reached when the conventional ATI analysis is replaced with such a partitioned analysis. They disconfirmed the findings of a previous study which had detected a significant ATI for the effects of instructional method on the aptitude-math achievement relationship (Cronbach & Webb). Cronbach and Webb's reanalysis of these data found that there was no evidence for a within-classes ATI and no conclusion could be reached about the between-classes analysis.

The within-context ATI is most commonly of interest when one wants to investigate the interaction effect of some type of class or school practice and aptitude on an outcome measure. In the example outlined previously the within-context interaction would be the difference in the home environment-achievement relationship between students receiving different levels of instructional technology, when controlling for the average home environment students in the school come from. The between-context model, on the other hand, would be the difference in the average home environment-achievement relationship among the different levels of instructional technology. Since scores are aggregated to the group level, the within-school information is lost. Additionally, aggregation bias often results in a dramatic increase in the correlation between variables (Robinson, 1950). Despite this limitation, the separate analyses approach used by Cronbach and Webb (1975) and detailed by Cronbach and Snow (1977) can be effectively employed to interpret ATI effects from hierarchically nested samples, particularly when intraclass correlations are low.

Multilevel Modeling

Another technique which has been used to make cross-level inferences in school-effects study is multilevel modeling. Multilevel modeling has particular merit when analyzing data which has high intraclass correlations due to the hierarchically nested structure of the data. When analyzing data with intraclass correlations using conventional regression analysis, the data are forced to fit a model that does not reflect how they were collected. Conversely, multilevel techniques draw strength from appropriately modeling the data at each level of the sampling design. In multilevel modeling a separate micro-level model is defined for each macro unit. In a school-effects study this would mean that within-school regression coefficients are modeled by school-level variables (De Leeuw & Kreft, 1986).

Random Coefficients Models

A particular type of multilevel model that is often used to make cross-level inferences is one in which the regression coefficients are not assumed to be constant for all contexts. In the multilevel analysis literature these models are alternately called random line models, slopes and intercepts as outcomes, and random coefficient models. These multilevel models circumvent a limitation of the ATI separate analyses approach, the assumption that the interaction is assumed to be homogeneous for all contexts within a particular group. However, logic would dictate that the ATI would be different across schools, because strategies and policies, as well as environmental factors, often vary from school to school. In random coefficient regression models, the parameters are allowed to differ over the different schools and are treated as a function of school characteristics and random but unique school variations that are assumed to be constant in the ATI model.

In addition to providing a more realistic model of the data, the random coefficients model is technically also an improvement over the conventional multiple regression model because it calculates the correct standard errors. Moreover, the random coefficients model improves the estimation of the parameters for the separate schools. An empirical Bayes estimation procedure is used to weight the regression coefficient estimates of each school by a reliability coefficient calculated for each school. This process is known as shrinkage because the estimates are "shrunk" toward the estimated group mean coefficients. Those schools providing less reliable estimates experience the most shrinkage (Cheung, Keeves, Sellin & Tsoi, 1990; Raudenbush, 1988). The resulting shrinkage estimates are more precise parameter estimates than those generated through ordinary least square methods.

The simplest random coefficients model is one in which there is an individual-level predictor (X), but no group-level predictors (Bryk & Raudenbush, 1992).

$$Y_{ij}=\beta_{0j}+\beta_{1j}(X_{ij}-\overline{X}_j)+r_{ij}$$
$$\beta_{0j}=\gamma_{00}+u_{0j}$$
$$\beta_{1j}=\gamma_{10}+u_{1j} \qquad (4)$$

In a school-effects study, the first equation is the micro-level model, in which a student-level outcome variable is a function of a student-level predictor, $(X_{ij}-\overline{X}_j)$, and the unique effect of the student, r_{ij}. The subsequent equations are the macro-level models which illustrate how the microparameters, β_0 and β_1 are modeled by school-level effects. β_{0j} is modeled as a function of the average outcome across the schools, γ_{00}, and the unique effect of school j on the mean outcome, m_{0j}. β_{1j} is modeled as a function of the average slope of the predictor-criterion relationship across the schools, γ_{10}, and the unique effect of each school j on the predictor-criterion slope, m_{1j}. This model can be extended to include other variables at both the micro-level, as well as the macro-level. Therefore, these models can be very useful for modeling cross-level inference. A random coefficients model that would appropriately partition the ATI into within- and between-context components is (Raudenbush, 1989):

$$Y_{ij}=\beta_{0j}+\beta_{1j}(X_{ij}-\overline{X}_j)+r_{ij}$$
$$\beta_{0j}=\gamma_{00}+\gamma_{01}\overline{X}_j+\gamma_{02}T_j+\gamma_{03}\overline{X}_jT_j+u_{0j}$$
$$\beta_{1j}=\gamma_{10}+\gamma_{11}T_j+u_{1j} \qquad (5)$$

In this multilevel model, γ_{03} is the between-contexts interaction, while γ_{11} is the within-contexts interaction[1]. Therefore, the ATI interaction can be partitioned into within- and between-contexts components, while still allowing for unique variation among schools. Other contextual variables can also be included in the macromodels to account for some of the variation unique to each school. For instance, school climate might be included in the example previously given to determine if the interaction between home environment and instructional technology is masked by school climate.

Recommendations and Conclusions

The ATI analysis method is suitable for making cross-level inferences when the intraclass correlation is low, and therefore the macro-level units are assumed to have a constant predictor-criterion relationship within each group. If there are any contextual effects, the partitioning of the interaction into within- and between-context components can aid in determining what factors are contributing to the interaction.

Hierarchical linear modeling is the recommended technique when intraclass correlations are high, and are not assumed to have a constant predictor-criterion relationship within each group. The main advantage of HLM techniques in this situation is that the parameters are assumed to be unique for each context and are modeled accordingly. Further, the shrinkage estimates of the parameters are more precise than those estimated through ordinary least squares analysis.

Whichever method is used to analyze the ATI in school-effects studies, one must keep in mind that most school-effects research is quasi-experimental. Students are not randomly assigned to schools, and schools are not often randomly assigned to treatment or practice. Therefore, confounding effects operating at the student level and the school

[1] Simplifying into one equation we arrive at $Y_{ij}=\gamma_{00}+\gamma_{01}\overline{X}_j+\gamma_{02}T_j+\gamma_{03}\overline{X}_jT_j+\gamma_{10}(X_{ij}-\overline{X}_j)+\gamma_{11}(X_{ij}-\overline{X}_j)T_j+u_{0j}+u_{1j}(X_{ij}-\overline{X}_j)+r_{ij}$

level may bias variance estimates and parameter estimates. This problem is exacerbated when one cannot define a set of reliable and valid measures to assess the school practice that is of interest (Raudenbush & Willms, 1995).

References

Bryk, A. S. & Raudenbush, S. W. (1992). *Hierarchical linear models: Applications and data analysis methods.* Newbury Park, CA: Sage Publications.

Cheung, K. C., Keeves, J. P., Sellin, N., & Tsoi, S. C. (1990). The analysis of multilevel data in educational research: Studies of problems and their solutions. *International Journal of Educational Research, 14,* 215-319.

Cronbach L. J. & Webb, N. (1975). Between-class and within-class effects in a reported aptitude x treatment interaction: Reanalysis of a study by G. L. Anderson. *Journal of Educational Psychology, 67,* 717-724.

Cronbach L. J. & Snow, R. E. (1977). *Aptitudes and instructional methods.* New York: Irvington.

De Leeuw, J. D. & Kreft, I. (1986). Random coefficient models for multilevel analysis. *Journal of Educational Statistics, 11,* 57-85.

Draper, D. (1995). Inference and hierarchical modeling in the social sciences. *Journal of Educational and Behavioral Statistics, 20,* 115-147.

Kreft, I. & Yoon, B. (1994). *Are multilevel techniques necessary? An attempt at demystification.* Presented at the annual meeting of the American Educational Research Association, New Orleans, LA.

Raudenbush, S. W. (1988). Educational applications of hierarchical linear models: A review. *Journal of Educational Statistics, 13,* 85-116.

Raudenbush, S. W. (1989). A response to Longford and Plewis. *Multilevel Modeling Newsletter, 1,* 8-10.

Raudenbush, S. W. & Bryk, A.. S. (1986). A hierarchical model for studying school effects. *Sociology of Education, 59,* 1-17.

Raudenbush, S. W. & Willms, J. D. (1995). The estimation of school effects. *Journal of Educational and Behavioral Statistics, 20,* 307-335.

Robinson, W. S. (1950). Ecological correlations and the behavior of individuals. *American Sociological Review, 15,* 351-357.

Missing Cells

continued from page 3

Williams, J.D. (1977a). Full rank and non-full rank models with contrast and binary coding systems for two-way disproportionate cell frequency analyses. *Multiple Linear Regression Viewpoints, 8*(1), 1-17.

Williams, J.D. (1977b). A note on coding the subjects effect in treatments x subjects designs. *Multiple Linear Regression Viewpoints, 8*(1), 32-35.

Williams, J.D. (1981). Missing cells in a two-way classification. *Multiple Linear Regression Viewpoints, 11*(1), 91-105.

Williams, J.D. (1991). Age, cohort and period in life-span research: A three-way analysis with logically missing cells. *Multivariate Behavioral Research, 26,* 631-654.

Williams, J.D., & Klug, M. (in press). Aging and cognition - Methodological differences in outcome. *Experimental Aging Research.*

Williams, J.D., & Wali, M. (1979). Missing cells and a curious case of degrees of freedom. *Multiple Linear Regression Viewpoints, 9*(5), 75-87.

Williams, J.D., & Williams, J.A. (1985). Testing hypotheses in a repeated measures design: An example. *Multiple Linear Regression Viewpoints, 13*(2), 36-46.