How to Apply Statistical Software in Doing Multilevel Modeling: A Comparative Study Perspective

Liming Lin\textsuperscript{1} and Xiaosan Zhang\textsuperscript{2}\textsuperscript{*}

\textsuperscript{1} School of Internet Economics and Business, Fujian University of Technology, Fuzhou 350018, China
\textsuperscript{2} Research Centre for Belt & Road Financial and Economic Development, Xiamen National Accounting Institute, Xiamen 361005, China
Email: 443918678@qq.com

Abstract. Multilevel model methodology is widely used in social science. There are many textbooks and papers introducing the principles and methods of multilevel model. However, most of them are simply a brief instruction how to apply statistical software to run multilevel model, which brings some obstacles for beginners in empirical research. To address this gap, this paper aims to help beginners and common users to grasp the operation skills of multilevel model in these statistical software. Firstly, the operation procedure of multilevel model in eight statistical software are introduced. Secondly, this paper systematically compares the output results of multilevel models of these statistical software. Finally, some operational guidelines and suggestions are put forward for the empirical application of multilevel model.

1. Introduction
Hierarchical data structure prevalently exists in research of social science [1] such as psychology, education, management, sociology, etc. Multilevel model, put forward by Lindley and Smith in 1972 [2], is applied to analyze hierarchical data. Although there are many textbooks and papers describing the statistical principles and methods of multilevel modeling [1, 3], less attention is paid to how to use software to run multilevel modeling. Some descriptions center on HLM software but rarely on others. With GUI (graphical user interface) and cumbersome process, it is difficult for beginners to use HLM to run multilevel modeling.

Statistical software which can be used to run multilevel models include HLM, LISREL, MLwiN, Mplus, R, SAS, SPSS, and Stata, but they are much different in doing multilevel model. Therefore, this paper initially introduces how to use the above 8 software for multilevel model analysis. Then, an example is given to systematically compare the outputs of a two-level linear model estimated by these software. Finally, this paper explores several key issues about how to build multilevel model application, and to provide some operational guidance and suggestions for doing empirical research.

2. Introduction of 8 Statistical Software for Multilevel Modeling
(1) HLM. HLM provides the GUI to do multilevel model and supports almost all data file formats. The default parameter estimation method includes REML (Restricted estimation maximum likelihood) and FIML (Full information maximum likelihood method) [4].

When using HLM, firstly click on “Make new MDM file” under “File” to choose “Stat package input” and “HLM2”; then move to MDM-HLM2”, choose “SPSS/Windows” under “Input File Type”, select “persons within groups” under “Nesting of Input data default”, and click on “Browse” to select the first
and second layers of data files under Level-1 and Level-2 Specification. Next click on “Choose Variables” to choose primary key, dependent variables, level-1 and level-2 independent variables. Then click on “Make MDM”, “Check Stats”, “Done” to enter into model settings interface to define dependent, independent and centralized variables. And set fixed or random effects in parameters estimation. Finally, click on “Run Analysis”, and the location and name of output file can be acquired in “Basic Settings”.

(2) LISREL. LISREL provides GUI or program commands for multilevel modeling [5]. ML (Maximum likelihood) is the only default parameters estimation method.

Firstly, click on “File” -> “New”, choose “PRELIS Data”, and import external files using “Import Data ...” under “File”. Secondly, click on “Multilevel” -> “Linear model” -> “Title and Options”, set the title, the number of iterations, convergence criteria, etc. Thirdly, set the Primary Key Variable to identify a group or hierarchical variable in Identification Variables window, then click on “Next”. Fourthly, set dependent and independent variables with fixed effects in “Select Response and Fixed Variables”, then click on “Next”. Finally, set the independent variables with random effects in the “Random Variables”, click on “Finish” to complete model settings. LISREL will generate .PR2 file where the model can be modified, click on the “P” to finish the model estimation, and then check .OUT file.

(3) MLwiN. MLwiN provides GUI for multilevel modeling. Its default estimation methods include IGLS (Iterative Generalized Least Square), RIGLS (Restricted IGLS), and MCMC (Markov Chain Monte Carlo) [6]. Firstly click on “File” -> “New worksheet” to generate .WS file, click on “File”->“Import” to import external data file, and then click on “Model” -> “Equation” to enter the model setting. Click on “Notation” to choose the subscript representation of the variable, and click on “y” to select the dependent variable and the number of model levels. Next click on “Add Terms” to add level-1 and level-2 independent variables, and add random effects in the model by clicking on the variables. Then, click on the “Start” to complete the model estimation. Finally click on the “Estimate” button below to get the results.

(4) Mplus. Mplus does not provide GUI but offer commands and programs to do multilevel model [7]. Mplus provides the most parameter estimation methods among these eight software, including ML, FIML, MLR, WLS and so on.

Mplus's program for multilevel model consists of title, data, variables, analysis, model, and output. Analysis type and parameter estimation method should be specified for analysis section, the model be set in model section, and the results be specified in output section.

(5) R. R does not provide GUI and multilevel models can only be done by writing program commands. The default parameter estimation method is REML [8]. Firstly import the data file, then use the “imer” function command to do multilevel model analysis, and finally use the summary function to get the results.

(6) SAS. SAS provided no GUI, and multilevel model analysis is implemented using the “MIXED” program command, and the parameter estimation method is REML by default [9]. Firstly import the data file, then uses the mixed program to set up the data and model, and the output is presented in the result viewer.

(7) SPSS. SPSS provides user-friendly GUI and programming commands. The default parameter estimation is REML or ML [10]. When data file is imported, click on “Analyze” -> “Mixed Models” -> “Linear”, import key variables into “Subjects”, and click on “Continue” to enter model settings window. Then import dependent variables into “Dependent Variable”, category variable into Factor (s), and continuous variable into Covariate (s). Next click on “Fixed” to set independent variable for fixed effect, then click on “Random” to set variable for random effect, and click on “Estimation” to set parameter estimation, click on “Statistics” to select output, click on “EM Means” to set estimation margin means. Finally, click on the “OK” to get model estimation result in the output.

(8) Stata. Stata perform multilevel model via program command with a simple and intuitive operation menu. The default parameter estimation has ML and REML [11].
Although multilevel model analysis can be done in the menu operations, most users prefer using command program to do the analysis. The command for multilevel model analysis is “mixed” and the basic syntax is as follows:

```
mixed depvar fe_equation [|| re_equation] [|| re_equation ...] [, options]
```

3. Model Specification Issues

This paper uses data of the American Tobacco Industry Association as an example to illustrate multilevel model with two-layer structure. This example aims to figure out factors influencing the voting behavior of U.S. congressman on bills involving the tobacco industry from 1997 to 2000. The data is of a two-level structure in which Level-1 is of congressman and Level-2 is of state where congressman come from. The dependent variable is the percentage of votes voted by members supporting tobacco industry. Level-1 independent variables includes political party (0 denotes Democrats and 1 denotes Republicans) and the amount of political donations congressman received from tobacco industry. Level-2 independent variable is the acres for cultivating tobacco in the state where congressman come from.

The two-level linear model includes several different conditions. The first basic model only includes a random intercept, then adds the fixed effect of level-1 explanatory variables to form the second model. The third model incorporates the random effect of level-1 explanatory variables based on the second model, and then adds level-2 explanatory variables with no interaction to form the fourth model. Finally, based on the fourth model, the fifth model adds the interaction effect of level-1 and level-2 explanatory variable. All these five models will be estimated by HLM, LISREL, MLwiN, Mplus, R, SAS, SPSS, and Stata separately. In addition, multilevel model has two forms of expression: one is hierarchical expression mode which is easier to distinguish different levels of structure while the other is mixed mode which focuses on the convenience of identifying random and fixed effects. The multilevel model described below is represented by hierarchical expression mode and mixed effect mode respectively.

Hofmann and Gavin (1998) finds that in cross-level interaction models, non-centralization is statistically equivalent to centralization in the model [12]. Since non-centralization and centralization with grand-mean differ only in intercept, and centralization with group-mean are not used in most cases, this paper employs all variables of original data into the model estimation.

There are many parameter estimation methods for multilevel modeling. MLwin software has IGLS and RIGLS, the default method in HLM, R, SAS, and SPSS uses REML, and Mplus and Stata use ML as default method. The estimation results of ML and REML are similar while the estimation results of IGLS, RIGLS and ML and REML are slightly different.

ICC (intra-group correlation coefficient) refers to the proportion accounted for by hierarchical model in variance of the dependent variable, determining whether observations in the sample data are independent, which in turn can illustrate the relevance of multilevel model application. The larger the ICC, the more necessary a multilevel model should be used. The ICC indicator is calculated as follows:

\[
ICC = \frac{\sigma^2_{u0}}{\sigma^2_{u0}+\sigma^2_e}
\]  

\(\sigma^2_{u0}\) denotes the variance of the residuals of level-2 regression equation. \(\sigma^2_e\) denotes the variance of the residuals of level-1 regression equation.

4. Comparison of Model Estimation Results

(1) Model 1: Random Intercept model (zero model)

Mixed model

\[
vote_{ij} = \gamma_{00} + u_{0j} + e_{ij}
\]  

Hierarchical model

\[
vote_{ij} = \beta_{0j} + e_{ij}
\]

\[\beta_{0j} = \gamma_{00} + u_{0j}\]
In model 1, $\gamma_{00}$ denotes the fixed effect of the intercept—the mean of dependent variable, $u_{0j}$ represents the random effect of intercept, and $e_{ij}$ is the error term. $\sigma^2_{u_{0j}}$ is the variance of $u_{0j}$, and $\sigma^2_{e_{ij}}$ denotes the variance of the model residual $e_{ij}$.

Table 1 shows that the estimation results of model 1 by HLM, LISREL, MLwiN, Mplus, R, SAS, SPSS, and Stata are almost the same. The estimation result of $\gamma_{00}$ is 0.528, the standard error is 0.031 or 0.032; the standard errors of $\sigma^2_{u_{0j}}$ and $\sigma^2_{e_{ij}}$ are almost equivalent. In particular, HLM and R do not report the standard errors of $\sigma^2_{u_{0j}}$ and $\sigma^2_{e_{ij}}$. The ICC of model 1 is $0.038 / (0.038+0.091) = 0.295$, which means that the factor of the congressman’s state accounts for nearly 30% of the variance of the percentage of the congressman’s vote in favor of the tobacco industry, indicating that it is necessary to incorporate the factors in the congressman’s state for multilevel model analysis.

Table 1. Estimation results of model 1.

| Software | Intercept $\gamma_{00}$ | HLM | LISREL | MLwiN | Mplus | R | SAS | SPSS | Stata |
|----------|--------------------------|-----|--------|-------|-------|---|-----|------|-------|
| **Fixed effects** | | 0.528 | 0.528 | 0.528 | 0.528 | 0.528 | 0.528 | 0.528 | 0.528 |
| (missing) | | (0.032) | (0.031) | (0.031) | (0.031) | (0.032) | (0.032) | (0.032) | (0.031) |
| **Random effect** | | $0.038$ | $0.037$ | $0.037$ | $0.037$ | $0.038$ | $0.038$ | $0.038$ | $0.037$ |
| (Intercept) | | (missing) | (0.010) | (0.010) | (0.009) | (missing) | (0.010) | (0.010) | (0.010) |
| $\sigma^2_{u_{0j}}$ | | $0.091$ | $0.091$ | $0.091$ | $0.091$ | $0.091$ | $0.091$ | $0.091$ | $0.091$ |
| (missing) | | (0.006) | (0.006) | (0.009) | (missing) | (0.006) | (0.006) | (0.006) | (0.006) |

(2) Model 2: Random Intercept with Fixed Level-1 Factor

Mixed model

$$votepct_{ij} = \gamma_{00} + \gamma_{10}party_{ij} + \gamma_{20}money_{ij} + u_{0j} + e_{ij}$$  \(5\)

Hierarchical model

$$votepct_{ij} = \beta_{0j} + \beta_{1j}party_{ij} + \beta_{2j}money_{ij} + e_{ij}$$  \(6\)

$$\beta_{0j} = \gamma_{00} + u_{0j}$$  \(7\)

$$\beta_{1j} = \gamma_{10}$$  \(8\)

$$\beta_{2j} = \gamma_{20}$$  \(9\)

Comparing to model 1, model 2 adds two level-1 independent variables. $\gamma_{10}$ denotes the fixed effect of the party of which congressmen are; $\gamma_{20}$ indicates the fixed effect of the political donations received by the congressmen.

Table 2 shows the estimation results of parameter given by these 8 software are almost the same. The estimation result of $\gamma_{00}$ is 0.209, the standard error is 0.019; the estimation result of the fixed effect of political parties $\gamma_{10}$ is 0.487, with the standard error 0.017 or 0.016; and the estimation result of the fixed effect of political donations $\gamma_{20}$ is 0.004 and the standard error is 0.000. The standard errors of $\sigma^2_{u_{0j}}$ and $\sigma^2_{e_{ij}}$ are almost the same. Specifically, HLM and R do not report the standard errors of $\sigma^2_{u_{0j}}$ and $\sigma^2_{e_{ij}}$. The ICC of model 2 is $0.009 / (0.009+0.031) = 0.225$, which indicates that the ICC dropped to 0.225 after adding these two fixed factors at the congressmen’s levels.

(3) Model 3: Random Intercept and Slope for Level-1 Factor

Mixed model

$$votepct_{ij} = \gamma_{00} + \gamma_{10}party_{ij} + \gamma_{20}money_{ij} + u_{0j} + u_{1j}party_{ij} + u_{2j}money_{ij} + e_{ij}$$  \(10\)
Table 2. Estimation results of model 2.

| Software    | HLM  | LISREL | MLwiN | Mplus | R   | SAS  | SPSS | Stata |
|-------------|------|--------|-------|-------|-----|------|------|-------|
| Intercept ($\gamma_{00}$) | 0.209 (0.019) | 0.209 (0.019) | 0.209 (0.019) | 0.209 (0.026) | 0.209 (0.019) | 0.209 (0.019) | 0.209 (0.019) | 0.209 (0.019) |
| Party ($\gamma_{10}$) | 0.487 (0.017) | 0.487 (0.016) | 0.487 (0.016) | 0.487 (0.025) | 0.487 (0.017) | 0.487 (0.017) | 0.487 (0.016) | 0.487 (0.016) |
| Money ($\gamma_{20}$) | 0.004 (0.000) | 0.004 (0.000) | 0.004 (0.000) | 0.004 (0.001) | 0.004 (0.000) | 0.004 (0.000) | 0.004 (0.000) | 0.004 (0.000) |
| $\sigma^2_{\text{Random effect}}$ | | | | | | | | |
| $\sigma^2_{\epsilon_{ij}}$ (Intercept) | 0.031 (missing) (0.002) | 0.031 (missing) (0.002) | 0.031 (missing) (0.003) | 0.031 (missing) (0.002) | 0.031 (missing) (0.002) | 0.031 (missing) (0.002) | 0.031 (missing) (0.002) | 0.031 (missing) (0.002) |

Hierarchical model

\[
\text{votept}_{ij} = \beta_{0j} + \beta_{1j} \text{party}_{ij} + \beta_{2j} \text{money}_{ij} + e_{ij}
\]  
(11)

\[
\beta_{0j} = \gamma_{00} + u_{0j}
\]  
(12)

\[
\beta_{1j} = \gamma_{10} + u_{1j}
\]  
(13)

\[
\beta_{2j} = \gamma_{20} + u_{2j}
\]  
(14)

Comparing to model 2, model 3 adds the random effect of the two level-1 independent variables. $u_{1j}$ denotes the random effect of the party factor of the congressmen, with the standard error $\sigma^2_{u_{1j}}$; $u_{2j}$ represents the random effect of the political donations received by the congressmen, with the standard error $\sigma^2_{u_{2j}}$.

The estimation results of each parameter in model 3 given by the HLM, LISREL, R, SAS, and SPSS are almost the same, and the estimation results from Mplus and Stata are the same, but MLwiN gives different results from other software. The estimation result of the intercept item $\gamma_{00}$ is 0.231, the standard error is 0.028; the estimation result of the fixed effect of political parties $\gamma_{10}$ is 0.468, with the standard error 0.023; and the estimation result of the fixed effect of political donations $\gamma_{20}$ is 0.005 and the standard error is 0.001. In particular, HLM and R do not report the standard errors of $\sigma^2_{u_{1j}}$ and $\sigma^2_{u_{2j}}$. The ICC of model 3 is 0.026/(0.026+0.028) = 0.481, which indicates that nearly 50% of the variance of congressmen who voted for the tobacco industry is explained by the factors in the state where the congressmen come from after adding these two fixed and random effect variables of the congressmen’s levels.

(4) Model 4: Random Level-1 Factors and Fixed Level-2 Factor (No Interactions)

Mixed model

\[
\text{votept}_{ij} = \gamma_{00} + \gamma_{01} \text{acres}_{i} + \gamma_{10} \text{party}_{ij} + \gamma_{20} \text{money}_{ij} + u_{1j} \text{party}_{ij} + u_{2j} \text{money}_{ij} + u_{0j} + e_{ij}
\]  
(15)

Hierarchical model

\[
\text{votept}_{ij} = \beta_{0j} + \beta_{1j} \text{party}_{ij} + \beta_{2j} \text{money}_{ij} + e_{ij}
\]  
(16)

\[
\beta_{0j} = \gamma_{00} + \gamma_{01} \text{acres}_{i} + u_{0j}
\]  
(17)

\[
\beta_{1j} = \gamma_{10} + u_{1j}
\]  
(18)

\[
\beta_{2j} = \gamma_{20} + u_{2j}
\]  
(19)
Comparing to model 3, model 4 adds a level-2 independent variable which is the acres for cultivating tobacco in the state where the congressman comes from. $\gamma_{01}$ denotes the fixed effect of level-2 independent variable which is the acres for cultivating tobacco in the state.

The estimation results of each parameter in model 4 given by the HLM, LISREL, R, SAS, and SPSS are almost the same, and the estimation results from Mplus, Stata are the same, but MLwiN gives different results from other software. The estimation result of the intercept item $\gamma_{00}$ is 0.220, the standard error is 0.025; the estimation result of the fixed effect of the acres for cultivating tobacco $\gamma_{01}$ is 0.001, with the standard error 0.000; the result for fixed effect of party $\gamma_{10}$ is 0.470, with the standard error 0.023; and the estimation result of the fixed effect of political donations $\gamma_{20}$ is 0.004, with the standard error is 0.001. Particularly, HLM and R do not report the standard errors of $\sigma_{u0j}$ and $\sigma_{eij}^2$. The ICC of model 4 is 0.020/ (0.020+0.027) =0.426, which indicates that after adding the fixed effect and the random effect of two level-1 independent variables and the fixed effect of one level-2 independent variable, 42.6% of the variance of congressmen who voted for the tobacco industry is explained by the factors in the state where the congressmen come from, which, compared with model 3, drop by 5.5% explained by the acres of cultivating tobacco in the state.

(5) Model 5: Random Level-1 Factors and Fixed Level-2 Factor with Interaction

Mixed model

\[ \text{votecpct}_{ij} = \gamma_{00} + \gamma_{01} \text{acres}_{ij} + \gamma_{10} \text{party}_{ij} + \gamma_{20} \text{money}_{ij} + \gamma_{11} \text{party}_{ij} \times \text{acres}_{ij} + \gamma_{21} \text{money}_{ij} \times \text{acres}_{ij} + u_{0j} \text{party}_{ij} + u_{2j} \text{money}_{ij} + u_{0j} + e_{ij} \]  

Hierarchical model

\[ \text{votecpct}_{ij} = \beta_{0j} + \beta_{1j} \text{party}_{ij} + \beta_{2j} \text{money}_{ij} + e_{ij} \]  

\[ \beta_{0j} = \gamma_{00} + \gamma_{01} \text{acres}_{ij} + u_{0j} \]  

\[ \beta_{1j} = \gamma_{10} + \gamma_{11} \text{acres}_{ij} + u_{1j} \]  

\[ \beta_{2j} = \gamma_{20} + \gamma_{21} \text{acres}_{ij} + u_{2j} \]  

Comparing to model 4, model 5 adds interactions between two level-1 independent variables and one level-2 independent variable. $\gamma_{11}$ denotes the interaction between party and acres and $\gamma_{21}$ denotes the interaction between political donations and acres.

The estimation results of each parameter in model 5 given by the HLM, LISREL, R, SAS, and SPSS are slightly different. Taking HLM as an example, the estimation result of the intercept item $\gamma_{00}$ is 0.183, the standard error is 0.021; the estimation result of the fixed effect of the acres for cultivating tobacco $\gamma_{01}$ is 0.003, with the standard error 0.001; the fixed effect of party $\gamma_{10}$ is 0.498, with the standard error 0.022; and the estimation result of the fixed effect of political donations $\gamma_{20}$ is 0.005 with the standard error is 0.001. The interaction between party and the acres is -0.002, with the standard error 0.0005 while the interaction between political donations and the acres is -3e-5, with the standard error 8e-6. The ICC of model 5 is 0.010/ (0.010+0.027) =0.270, which indicates that after adding fixed effect and random effect of two level-1 independent variables, fixed effect of one level-2 independent variable, and interactions between two level-1 independent variables and one level-2 independent variable, 27.0% of the variance of congressmen who voted for the tobacco industry is explained by the factors in the state where the congressmen come from, which, compared with model 4, drop by 15.5% explained by the interactive effect.

5. Discussion

According to the empirical results above, we try to give some suggestions to the researchers in multilevel modeling. For beginners, it is more convenient to use statistical software with GUI, such as HLM, SPSS, and MLwiN. The operational procedure of HLM is more complicated, but it would be more helpful to understand the method of multilevel modeling. Additionally, it is relatively cumbersome to use SPSS and difficult to understand the process of multilevel modeling, but since many questionnaire researchers are used to applying SPSS software, it is still worth learning to use SPSS to perform multilevel modeling.
If researchers have a good background at using these statistical software, we suggest them to use software with multilevel modeling commands, such as R, SAS, Stata, especially R to obtain estimation results more quickly.

6. Conclusion
This paper briefly describes the characteristics and the application areas of the multilevel modeling. Taking the simplest zero model as an example, we elaborate the operating procedure and code structure of the eight statistical software, and systematically depicts the estimation methods of the multilevel modeling. Then, based on five common multilevel models, and using the data of the American Tobacco Industry Act from 1997 to 2000, this paper compares the estimation results for these five multilevel models run by these eight software above. Finally, we try to offer some guidelines in doing multilevel analysis, and hope it would be useful to researchers in this field.

Reference
[1] Snijders T A B and Bosker R J 2011 *Multilevel analysis: An Introduction to Basic and Advanced Multilevel Modeling* (London: Sage).
[2] Lindley D V and Smith A F M 1972 Bayes estimates for the linear model *J. R. Stat. Soc. B* 34 (1) 1-41.
[3] Raudenbush S W and Bryk A S 2002 *Hierarchical Linear Models: Applications and Data Analysis Methods* (London: Sage).
[4] Raudenbush S W, Cheong Y F and Congdon R T 2019 *HLM 8 for Windows* (Skokie, IL: Scientific Software International, Inc.).
[5] Jöreskog K G and Sörbom D 2018 *LISREL 10 for Windows* (Skokie, IL: Scientific Software International, Inc.).
[6] Wright D B 2011 *Statistical Software Review: MLwiN 2.0*. BRIT J. Math. Stat. Psy. 57 (2) 383-384.
[7] Muthén B and Asparouhov T 2011 Beyond multilevel regression modeling: Multilevel analysis in a general latent variable framework *Handbook of Advanced Multilevel Analysis* (New York: Taylor and Francis) Hox J and Roberts J K eds pp 15-40.
[8] Bates D, Mächler M and Bolker B 2015 Fitting linear mixed-effects models using lme4 *J. Stat. Softw.* 67 (1) 133-199.
[9] Singer J D 1998 Using SAS PROC MIXED to fit multilevel models, hierarchical models, and individual growth models *J. Educ. Behav. Stat.* 23 (4) 323-355.
[10] Heck R H, Thomas S L and Tabata L N 2010 *Multilevel and Longitudinal Modeling with IBM SPSS* (New York: Routledge).
[11] Sophia R-H and Anders S 2012 *Multilevel and Longitudinal Modeling Using Stata* 3rd ed (Texas: StataCorp LP).
[12] Hofmann D A and Gavin M B 1998 Centering decisions in hierarchical linear models: Implications for research in organizations *J. Manage.* 24 (5) 623-641.