DOCUMENT RESUME

ED 464 931 TM 033 861

AUTHOR	Miyazaki, Yasuo
TITLE	Johnson-Neyman Type Technique in Hierarchical Linear Model.
PUB DATE	2002-04-00
NOTE	66p.; Paper presented at the Annual Meeting of the American Educational Research Association (New Orleans, LA, April 1-5, 2002).
PUB TYPE	Reports - Descriptive (141) -- Speeches/Meeting Papers (150)
EDRS PRICE	MF01/PC03 Plus Postage.
DESCRIPTORS	Mathematical Models; Prediction; *Regression (Statistics)
IDENTIFIERS	*Hierarchical Linear Modeling; *Johnson Neyman Technique; Mathematica

ABSTRACT

One of the innovative approaches in the use of hierarchical linear models (HLM) is to use HLM for Slopes as Outcomes models. This implies that the researcher considers that the regression slopes vary from cluster to cluster randomly as well as systematically with certain covariates at the cluster level. Among the covariates, group indicator variables at the cluster level, which classify the cluster units into several groups, are often found to be significant predictors. If this is the case, the average relationships between the outcome and a key independent variable are different from group to group. Then the question arises, "At what range of the independent variable is the outcome statistically significantly different between groups?" The Johnson-Neyman (J-N) technique answers this kind of question in analysis of covariance (ANCOVA) settings. In the multi-level modeling context, the F-test, which is used in ANCOVA, cannot be applied because the assumption of homogeneity of variance within cluster units is violated in most cases of data that have multi-level structure. Instead, the approximate Walt test can be used to determine the region of significance. The Mathematica computer software package, which is capable of symbolic processing, leads to a direct solution. Two examples from education and child development are provided in order to illustrate the technique and to show how to implement it with Mathematica. (Contains 10 figures, 6 tables, and 16 references.) (Author/SLD)

Johnson-Neyman Type Technique in Hierarchical Linear Model

Yasuo Miyazaki

This paper is prepared for the:
Annual Meeting of the American Educational Research Association in New Orleans, LA
April 2002

Abstract

One of the innovative insights in the hierarchical linear models (HLM) is to use it for "Slopes-as-Outcomes" models. This implies that we consider that the regression slopes vary from cluster to cluster randomly as well as systematically with certain covariates at the cluster level. Among the covariates, group indicator variables at the cluster level, which classify the cluster units into several groups, are often found to be significant predictors. If this is the case, the average relationships between the outcome and a key independent variable are different from group to group. Then, a question such as "at what range of the independent variable the outcome is statistically significantly different between groups?" naturally arises. Johnson-Neyman (J-N) technique answers this kind of question in the analysis of covariance (ANCOVA) settings. In multilevel modeling context, F-test, which is used in ANCOVA, can not be applied because the assumption of homogeneity of variance within cluster units is violated in most cases of the data that have multilevel structure. Instead, the approximate Walt test can be used to determine the region of significance. Mathematica computer software package which is capable of symbolic processing allows us to directly obtain the solution. Two examples from education and child development are provided in order to illustrate the technique and to show how to implement it by Mathematica.

In hierarchical linear models (HLM), it is conceived that each macro-unit has its own regression parameters and that those parameters vary randomly from macro-unit to macro-unit. This implies that each macro-unit has different regression slopes. At macro-level, we try to explain why there is variability of the dependent variable among macro-units using the macro-level independent variables, i.e., information about the characteristics belonged to the macro units. Frequently, the information includes qualitative variables which creates groups. For example, in the school-effectiveness study such as the High School and Beyond Survey (Coleman, Hoffer, & Kilgore, 1982) where the data have structure that students are nested within school, suppose one want to study the relationship between student's socioeconomic status (SES) and achievement. If the schools are grouped by the sectors such as private vs. public schools, then each sector has its own average relationship. If these intercepts and slopes are different, then it would be of interest to ask in which region of student's SES in fact the two sectors have significant difference.

In growth modeling via hierarchical modeling framework, repeated measures are conceived as nested within subjects and thus each subject has his/her own growth trajectory characterized by a set of random regression coefficients at level-1. Thus, those random coefficients vary from subject to subject. At level-2, those variability are tried to explain by the individual characteristics. Frequently, among those individual characteristics categorical variables such as gender, race/ethnicity that characterizes individual are included (For example, Huttenlocher et al, 1991; Raudenbush & Chan, 1993). In this case, we can obtain the average growth trajectory for each group. Then, a question such as "In which age are there statistically significant group differences in terms of expected values of the outcome?" can be interesting developmentally.

Thus, there are many instances that we are interested in the range of the covariates where the expected outcome has significant differences among groups in multilevel modeling. This is a natural consequence of multilevel modeling because it conceives that each macro unit has its own regression lines. And if there are the grouping variables that interact with independent variables, then, a question on which region of independent variable is actually statistically significant naturally arises. This question can be answered by the Johnson-Neyman procedure in Analysis of covariance (ANCOVA) context if the linear model is appropriate. Clearly, multilevel model does not satisfy the i.i.d. error assumption, the J-N technique can not be applied directly. However, the spirit of J-N technique can be applied to the multilevel modeling by changing the test statistic to be computed and the subsequently the reference distribution. That is, by casting the J-N problem into a linear hypothesis, we can use the Wald statistic in stead of F statistic. Though it is not exact as in the F-test in ANCOVA-linear model context, the Wald test is a good approximation when the sample size for the cluster units is large enough, as freqently the case for the settings to apply the multilevel models.

Johnson—Neyman Procedure in ANCOVA

The Johnson-Neyman technique, as it was originally formulated by Johnson and Neyman (1936), solves the problems of identifying regions of significance of the covariates in analysis of covariance (ANCOVA) when the regression lines are not parallel. The idea behind the J-N procedure can be described as follows. If we know the points of the covariates, we can test whether the means for several groups at the points are significantly different by constructing a F-statistic and then comparing the value to the corresponding critical value of the F-distribution because the test statistic is distributed as Fisher's F distribution if specification of the error distribution, which is specified as normal, is correct. The decision rule of rejecting the null hypothesis, such as group means are not different at the specified value of the covariates, provides the inequality that the value of the test statistic must satisfy in order to reject the null hypothesis. Since we don't know and actually we wish to know the points of the covariates, we make the values of the covariates unknown x and then solve the inequality with respect to the unknown value of the covariate, x. For the simple design such as two groups, i.e., treatment and control groups, and one covarites, a simple formula which basically solves a quadratic equation is appealing because it provides the insight on how the terms in the formula influence the solution. A relatively simple formula for the explicit solution is still available for several groups but only when the number of the covariates is one (Huitema, 1980, Chap. 13). If the number of covariates gets two, the formula gets complicated (Johnson and Fay, 1950), and if it is more than three, the formula is almost intractable.

Recently, the advent of computational softwares such as Mathematica (Wolfram, 1999) which explicitly provide the symbolic processing capabilities changes the formulation of Johnson-Neyman technique. That is, it can be phrased by general linear model framework and the general from of the equation does not have any limitation in terms of the number of covariates (Hunka, 1995; Hunka & Leighton, 1997). The disadvantage for this alternative formulation of the J-N technique was that it does not provide the immediate solution. Mathematica can overcome this limitation.

In the linear model, the dependent variable is regressed on the set of independent variables which include the qualitative grouping variables and the quantitative covariates. Suppose we have *g* groups and *p* covariates and we formulate the linear model with intercept. Then we have $(g\text{-}1)p$ independent variables and 1 intercept. In equation, the general linear model can be written as

$$Y = X\beta + \varepsilon\,, \tag{1}$$

where Y is the $n \times 1$ vector of observations, and X is a $n \times P$ design matrix where $P = (g-1)p + 1\, p \text{ out of } P$ are continuous covariates and β is the $P \times 1$ vector of parameters, and ε is the $n \times 1$ vector of errors and $\varepsilon \sim N(\mathbf{0}, \sigma^2 \mathbf{I})$ where $\mathbf{I}$ is the $n \times n$ identity matrix.

Whether the group differences are statistically significant at certain point of covariates, we assign the value for the covariates at the point we want to test and create contrast matrix K^T of order $df_k \times P$ where $rank(K^T) = df_k$ and the hypothesis we test can be expressed as:

$$H_0 : K^T \beta = \mathbf{0}\,. \tag{2}$$

The degrees of freedom for the contrast matrix K^T is the number of rows which are constructed as independent and is $df_k = g-1$ if we want to compare for all the groups.

To test the hypothesis, we use the statistic $\dfrac{SS_k / df_k}{SS_e / df_e}$, which will be distributed as F-distribution with the numerator degrees of freedom of df_k and the denominator degrees of freedom of df_e. The quantity SS_k in the numerator is the sum of squares of the contrast and is expressed as:

$$SS_k = (K^T\hat{\beta})^T[K^T(X^TX)^{-1}K]^{-1}K^T\hat{\beta} \tag{3}$$

where $\hat{\beta}$ is the least square estimator of β and it is $\hat{\beta} = (X^TX)^{-1}X^TY$. The quantity SS_e is the sums of squares of residual and it is written as $SS_e = (Y-\hat{Y})^T(Y-\hat{Y})$ where $\hat{Y}$ is the predicted value and is written as $\hat{Y} = X\hat{\beta}$. The mean squared error $MS_e = \dfrac{SS_e}{df_e}$ is the estimator of σ^2. The use of F distribution is justified because $\dfrac{SS_k}{\sigma^2} \sim \chi^2_{df_k}$ under H_0 and $\dfrac{SS_e}{\sigma^2} \sim \chi^2_{df_e}$ and SS_k and SS_e are independent.

Thus, the quantity $\dfrac{SS_k/df_k}{SS_e/df_e}$ is compared to the critical value $F_{\alpha;df_k;df_e}$ of the α-level test:

$$\frac{SS_k/df_k}{SS_e/df_e} \geq F_{\alpha;df_k;df_e}. \tag{4}$$

Therefore, for the test of hypothesis implied by K^T to be significant at α level the following inequality must hold:

$$(K^T\hat{\beta})^T[K^T(X^TX)^{-1}K]^{-1}K^T\hat{\beta} - (MS_e)(df_k F_{\alpha;df_k;df_e}) \geq 0. \tag{5}$$

The contrast matrix K^T includes the unknown values of covariates. The region of significance can be given by solving the above inequality with respect to the unknown values. Thus, to obtain the root of the equation directly by solving the equation

$$(K^T\hat{\beta})^T[K^T(X^TX)^{-1}K]^{-1}K^T\hat{\beta} - (MS_e)(df_k F_{\alpha;df_k;df_e}) = 0, \tag{6}$$

Mathematica is necessary since the computation requires symbolic algebra. For the case of two groups and one covariate, $(K^T\hat{\beta})^T[K^T(X^TX)^{-1}K]^{-1}K^T\hat{\beta}$ will be ratio of two quadratic polynomials.

If we want to have simultaneous regions of significance, which can be defined as a region such that, with confidence 1- α by a α-level test we can state that the groups are different simultaneously for all points contained in it for all possible pair of groups (Potthoff, 1964), we just replace the constant $df_k F_{\alpha;df_k;df_e}$ in the above inequality (2) by $(p+1)(g-1)F_{\alpha;(p+1)(g-1);df_e}$ where p is the number of predictor variables. The between non-simultaneous and simultaneous region of significance is that the constant part $df_k F_{\alpha;df_k;df_e}$ has larger value when we use the simultaneous region than the non-simultaneous region of significance. For example, when $g = 2$ and $p = 1$ (this is typical ANCOVA design where there are control and treatment group and there is a single covariate), we use $F_{\alpha;1;df_e}$ for non-simultaneous case and $2F_{\alpha;2;df_e}$ for the simultaneous case.

Note that we could use a chi-square test based on the fact that the quantity $\dfrac{SS_k}{MS_e}$ is approximately distributed as a chi-square with degrees of freedom df_k, and the decision rule for testing $H_0 : K^T\beta = \mathbf{0}$ in (2) is:

$$\text{Reject } H_0 \text{ if } \frac{SS_k}{MS_e} \geq \chi^2_{df=df_k,\alpha} \tag{7}$$

where $\chi^2_{df=df_k,\alpha}$ is the critical value for the α level test for the chi-square distribution with the degrees of freedom df_k. The justification for this can be seen by reminding that

$\frac{SS_k}{\sigma^2} \sim \chi^2_{df_k}$ and the MLE for σ^2 is MS_e. However, note that the nice thing about a test based on the F-distribution, unlike the test based on the chi-square distribution just mentioned, is that there is no need to know the true value of σ^2 because it cancels out in numerator and denominator. Another advantage of the F-test is that this is the exact test for any sample size as long as the model is correct. This creates the contrast with the approximate chi-square test, which is only exact for the large samples because the test relies on the asymptote of MS_e.

Johnson-Neyman type procedure in Hierarchical Linear Model

As I have described in the context of school effectiveness study and growth modelings, there are cases where the Johnson-Neyman type technique is required in multilevel model. That is, when the cross-level interaction exists, i.e., interaction between macro-level group membership and micro-level covariates[1], we pose a question: In which region of the covarites which group has the statistically significantly higher mean?

As in the J-N procedure for the linear model, the J-N type procedure in HLM can be expressed as the linear hypothesis on the fixed effects. Since the linear hypothesis testing is concerned about the fixed effects, the number of levels in the multilevel model doesn't matter. That is, we can apply the same procedure to the three level or higher hierarchical models. Here for the purpose of clarity of exposition, I present the two-level HLM model, the simplest case.

[1] Of course, the interactions which occur at the same level can be handled by the same way. But the cross-level interaction is mentioned here because it is a feature of multilevel modeling and the example scenarios were described by this cross-level interaction.

A two-level HLM can be written as:

Let Y_1, …, Y_j,.., Y_J are all independent and

$$Y_j = A_{1j}\gamma + A_{2j}u_j + \varepsilon_j, (j = 1,2,\ldots,J) \tag{8}$$

where Y_j is the $n_j \times 1$ vector of dependent variable; A_{1j} is the $n_j \times F$ matrix; A_{2j} is the $n_j \times R$ matrix; γ is the $F \times 1$ vector of fixed effects parameters; u_j is the $R \times 1$ vector of level-2 random effects; and ε_j is the $n_j \times 1$ vector of level-1 random effects. The distribution of u_j and ε_j are assumed to be $u_j \sim N_R(\mathbf{0}, \tau)$ where τ is a $R \times R$ positive definite symmetric matrix; and $\varepsilon_j \sim N(\mathbf{0}, \sigma^2 \mathbf{I}_{n_j})$ where σ^2 ia the level-1 error variance and $\mathbf{I}_{n_j}$ is the $n_j \times n_j$ identity matrix. The design matrix A_{1j} for the fixed effects parameter γ involves the qualitative grouping variable at level-2 and the covariates either at level-1 or level-2.

In ANCOVA we can use F test. But in HLM we need to use the Wald test (Wald, 1943) to test about the hypothesis on the fixed effects because in HLM the errors are not i.i.d., thus we cannot construct F. Since we use the Wald test, the test would be asymptotically correct when J goes to infinity.

Thus, the statistical test for Johnson-Neymann type hypothesis in HLM is a Wald test and has general form of linear hypothesis. Suppose we wish to test d_K linearly independent hypothesis on the fixed effect γ. Then using a $d_K \times F$ contrast matrix K^T, the null hypothesis can be written as

$$H_0\text{: } K^T\gamma = 0 \tag{9}$$

where $rank(K^T) = d_K$. And the statistic for testing the hypothesis (9) is

$$H = (K^T\hat{\gamma})^T \hat{V}_K^{-1}(K^T\hat{\gamma}) \tag{10}$$

where $\hat{V}_K^{-1}$ is the estimator of V_K^{-1} and

$$V_K \equiv Var(K^T\hat{\gamma}) = K^T Var(\hat{\gamma})K\,. \tag{11}$$

$\hat{\gamma}$ is the MLE of γ which is

$$\hat{\gamma} = (\sum_{j=1}^{J} A_{1j}^T V_j^{-1} A_{1j})^{-1} \sum_{j=1}^{J} A_{1j}^T V_j^{-1} Y_j \tag{12}$$

where

$$V_j = A_{2j}\tau A_{2j}^T + \sigma^2 \mathbf{I}_{\mathbf{n}_j}\,. \tag{13}$$

Then, the variance of $\hat{\gamma}$ is

$$Var(\hat{\gamma}) = (\sum_{j=1}^{J} A_{1j}^T V_j^{-1} A_{1j})^{-1}\,. \tag{14}$$

The MLE of $Var(\hat{\gamma})$ is obtained by substituting τ and σ^2 in V_j (see (13)) by their respective MLE. We denote it $\hat{Var}(\hat{\gamma})$. Thus $\hat{V}_K = K^T \hat{Var}(\hat{\gamma})K$. Note that $Var(\hat{\gamma})$ represented in Equation (14) is identical to the Fisher information obtained from taking the inverse of negative of the expectation of the second derivative of the log-likelihood[2], and thus $\hat{Var}(\hat{\gamma})$ is the fisher information evaluated at the MLE of σ^2 and

[2] The log-likelihood for the model (8) is

$l = \log[f(Y \mid \sigma^2, \tau, \gamma)] = -\frac{\sum_{j=1}^{N} n_j}{2}\log(2\pi) - \frac{1}{2}\sum_{j=1}^{J}\log|V_j| - \frac{1}{2}\sum_{j=1}^{J}(Y_j - A_{1j}\gamma)^T V_j^{-1}(Y_j - A_{1j}\gamma)$ and thus the information with respect to γ is $I(\gamma) = -E[\partial^2 l / \partial\gamma^2] = \sum_{j=1}^{J} A_{1j}^T V_j^{-1} A_{1j}$. Then, as the sample size increases, the MLE of γ is consistent and asymptotically normally distributed with mean γ and variance $[I(\gamma)]^{-1}$, which is represented by writing $\hat{\gamma} \sim AN(\gamma, [I(\gamma)]^{-1})$.

τ. Then the Wald statistic (Wald, 1941) H is constructed as $H = (K\hat{\gamma})^T \hat{V}_K^{-1}(K^T\hat{\gamma})$ as in (10), and the H will be distributed asymptotically, i.e., $J \to \infty$, as chi-square with the degree of freedom with the number of row in K^T which was denoted as n_K. Thus the statistic H is compared to the chi-square critical value to test whether the null hypothesis H_0 can be rejected at the specified α-level. Explicitly, we would reject H_0: $K^T\gamma = 0$ in (9) if

$$(K^T\hat{\gamma})^T \hat{V}_K^{-1}(K^T\hat{\gamma}) \geq \chi^2_{df=d_K;\alpha}. \tag{15}$$

To obtain the region of significance of the covariates, we use the following inequality

$$(K^T\hat{\gamma})^T \hat{V}_K^{-1}(K^T\hat{\gamma}) - \chi^2_{df=d_K;\alpha} \geq 0. \tag{16}$$

and solve for the unknown values in K^T.

It should be noted that the Wald test chosen for the J-N type technique in multilevel models corresponds to the alternative J-N technique in linear models which is represented by the decision rule (7). This method was also approximate and utilized a statistic that is approximately distributed as the chi-square distribution.

Illustrative Examples

Example 1. High school & Beyond Survey

The first example is taken from the High School & Beyond data set. The model seeks the relationship between mathematics achievement and the student's socio-economic status. The model was used in Bryk and Raudenbush (1992).

Briefly speaking, the data are collected in 1984 and this data is the U.S. representative sample. There are 7185 students in 160 schools. Out of 160, 70 schools are

Catholic and 90 schools are public high schools. The target variable of our interest is the student's mathematics achievement score and the mean for the students in public schools is 11.4 and that for Catholic schools is 14.2. The difference is 2.8, which is statistically significant. This difference is translated into 0.42 units of within-sector standard deviation unit (the pooled standard deviation (s.d.) = 6.734). The individual student's SES ranges from –3.76 to 2.69 with the mean of 0 and the standard deviation is 0.78, and the school mean SES ranges from –1.19 to .83 with the mean of 0 and the standard deviation of .41. Thus those variables have already been centered. In terms of sector difference, in Catholic schools the student's SES ranges from –3.76 to 2.69 with the mean of -0.15 and the standard deviation of 0.79, and in public schools, it ranges from –2.84 to 1.76 with the mean of 0.15 and the standard deviation of 0.74. Thus, on average the SES is 0.38 (the pooled s.d. = 0.78) higher in Catholic schools than public schools. For school SES, it ranges from –1.19 to 0.83 with the mean of 0 and the standard deviation of 0.41. Thus, this variable is also already centered. In terms of group differences based on the sector, the school SES ranges from –1.19 to 0.69 in public schools with the mean of –0.13 and the standard deviation of 0.38, and –0.76 to 0.83 in Catholic schools with the mean of 0.17 and the standard deviation of 0.40. Thus also in school SES, which can be conceived as a proxy of school resources, Catholic schools are more affluent by 0.76 standard deviation unit (0.389 for the pooled s.d.) than public schools.

To summarize, the effect size of the sector on Math achievement is 0.42, 0.38 on student's individual SES, and 0.76 on school SES. According to Cohen's definition of size of the effect size, the first two is considered to be in between small to medium, and the third one, the effect size of sector on school SES is considered to be large in social science research (Cohen, 1988). Students in Catholic schools do better in mathematics achievements than those in public schools; their parents are more educated and richer; and the schools are more resourceful. These descriptive statistics are summarized in Table 1.

Insert Table 1 About Here

The model to describe the relationship between mathematics achievement and SES for each sector, that was used in Bryk and Raudenbush (1992), is at level-1,

Model

Level-1:

$$Y_{ij} = \beta_{0j} + \beta_{1j}(SES_{ij} - \overline{SES}_{.j}) + \varepsilon_{ij}, \varepsilon_{ij} \overset{iid}{\sim} N(0, \sigma^2). \quad (17)$$

Thus at level-1, the model seeks the relationship between mathematics achievement Y_{ij} of the student i in school j and the his/her SES relative to the school mean SES. At level-2, the intercept β_{0j} and the SES slope β_{1j} vary from school to school depending on the school SES and the sector (public or Catholic), and those have random components, u_{0j} and u_{1j}:

Level-2:

$$\beta_{0j} = \gamma_{00} + \gamma_{01}(\overline{SES}_{.j} - \overline{SES}_{..}) + \gamma_{02}SECTOR_j + u_{0j}$$

(18)

$$\beta_{1j} = \gamma_{10} + \gamma_{11}(\overline{SES}_{.j} - \overline{SES}_{..}) + \gamma_{12} SECTOR_j + u_{1j},$$

where

$\begin{pmatrix} u_{0j} \\ u_{1j} \end{pmatrix} \overset{iid}{\sim} N(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} \tau_{00} & \tau_{01} \\ \tau_{10} & \tau_{11} \end{pmatrix}$; and $SECTOR_j = 1$ if the school j is Catholic and 0 if it is public; SES_{ij} is a measure of socio-economic status for student i in school j. It should be noted that at level-1 student's SES was centered around the school (group) mean and at level-2 school SES was centered around the grand mean. This specification decomposes the SES effect on mathematics achievement directly into within-school and between-school portions and thus makes the meanings of the parameters clear in terms of contextual effects. Also, the group-mean centering specification of the model at level-1 focuses on the relationship between Mathematics achievement and the students SES relative to the school mean SES. Thus, this specification creates that the regression surface depends on not only student's SES but also the school mean SES of the school to which the student belongs. The school SES, which is computed as the mean of the students' SES in the school, is sometimes interpreted as a proxy measure of the school's resources such as the number of teachers available, facilities, funding, and so forth.

The above model was fitted by the HLM software (Raudenbush, Bryk, Cheong, and Congdon, 2000) to the data via Restricted Maximum Likelihood Method (REML). The results are in the Table 2.

Insert Table 2 About Here

Note that in Table 2, the school SES was named as MEANSES.

$\hat{\gamma}_{00}$ = 12.10 indicates the average predicted score for the students in public school whose SES is the school mean and the MEANSES is the grand mean; $\hat{\gamma}_{10}$ = 2.94 indicates that on average one unit increase in student's SES increases the math achievement by 2.94 controlling for SECTOR and MEANSES. In terms of the sector effects, Catholic schools do better by 1.23 ($\hat{\gamma}_{02}$ = 1.23) when SES and MEANSES are controlled, but have lower SES slope ($\hat{\gamma}_{12}$ = -1.64). Thus, Catholic schools do better in excellence and provide more egalitarian education. MEANSES works in a way to increase both the intercept ($\hat{\gamma}_{01}$ = 5.33), i.e., the mean score given student's SES and SECTOR, and the SES slope ($\hat{\gamma}_{11}$ = 1.03). Since the mean mathematics achievement gap between Catholic and public schools decreases as the individual SES increases, we expect that from some point of individual SES to above, the public schools do better than Catholic schools. That is, the SES slopes are different for Catholic schools and public schools. In this context, naturally a question such as "In what range of SES, private and public schools are statistically different and which sector does better in that range?" arises and this question is especially important for prospective students and the parents who seek the best possible outcomes. This is the same situation that Johnson-Neyman procedure was called for.

In order to investigate the above question, we first need to know the expected regression planes for Catholic and public schools. Those are obtained by taking expectations over the distributions of both ε_{ij} and u_j;

$$\begin{aligned} E(Y_{ij}) &= \{\gamma_{00} + \gamma_{01}(\overline{SES}_{.j} - \overline{SES}_{..}) + \gamma_{02}(SECTOR)_j\} \\ &+ \{\gamma_{10} + \gamma_{11}(\overline{SES}_{.j} - \overline{SES}_{..}) + \gamma_{12}(SECTOR)_j\}(SES_{ij} - \overline{SES}_{.j}). \end{aligned} \tag{19}$$

Then, the expected achievements for Catholic and public sector are:

Catholic:

$$E(Y_{ij})_C = \{(\gamma_{00} + \gamma_{02}) + \gamma_{01}(\overline{SES}_{.j} - \overline{SES}_{..})\} + \{(\gamma_{10} + \gamma_{12}) + \gamma_{11}(\overline{SES}_{.j} - \overline{SES}_{..})\}(SES_{ij} - \overline{SES}_{.j}). \quad (20)$$

Public:

$$E(Y_{ij})_P = \{\gamma_{00} + \gamma_{01}(\overline{SES}_{.j} - \overline{SES}_{..})\} + \{\gamma_{10} + \gamma_{11}(\overline{SES}_{.j} - \overline{SES}_{..})\}(SES_{ij} - \overline{SES}_{.j}). \quad (21)$$

The fact that $\overline{SES}_{..} = 0$ simplifies (20) and (21) to:

Catholic:

$$E(Y_{ij})_C = \{(\gamma_{00} + \gamma_{02}) + \gamma_{01}\overline{SES}_{.j}\} + \{(\gamma_{10} + \gamma_{12}) + \gamma_{11}\overline{SES}_{.j}\}(SES_{ij} - \overline{SES}_{.j}). \quad (22)$$

Public:

$$E(Y_{ij})_P = \{\gamma_{00} + \gamma_{01}\overline{SES}_{.j}\} + \{\gamma_{10} + \gamma_{11}\overline{SES}_{.j}\}(SES_{ij} - \overline{SES}_{.j}). \quad (23)$$

Replacing the parameters to the estimates provides the predicted values as the function of SES_{ij} and $\overline{SES}_{.j}$.

Catholic:

$$\begin{aligned}\hat{Y}_{ij,C} &= \{(\hat{\gamma}_{00} + \hat{\gamma}_{02}) + \hat{\gamma}_{01}\overline{SES}_{.j}\} + \{(\hat{\gamma}_{10} + \hat{\gamma}_{12}) + \hat{\gamma}_{11}\overline{SES}_{.j}\}(SES_{ij} - \overline{SES}_{.j}) \\ &= (13.32 + 5.33\overline{SES}_{.j}) + (1.30 + 1.03\overline{SES}_{.j})(SES_{ij} - \overline{SES}_{.j}).\end{aligned} \quad (24)$$

Public:

$$\begin{aligned}\hat{Y}_{ij,P} &= \{\hat{\gamma}_{00} + \hat{\gamma}_{01}\overline{SES}_{.j}\} + \{\hat{\gamma}_{10} + \hat{\gamma}_{11}\overline{SES}_{.j}\}(SES_{ij} - \overline{SES}_{.j}) \\ &= (12.10 + 5.33\overline{SES}_{.j}) + (2.94 + 1.03\overline{SES}_{.j})(SES_{ij} - \overline{SES}_{.j})\end{aligned} \quad (25)$$

The predicted values are on the surface of the student's SES (SES_{ij}) and the school mean SES ($\overline{SES}_{.j}$) axes. We see several sections of the surface by fixing either SES_{ij} or $\overline{SES}_{.j}$ at three different values, i.e., one standard deviation above the mean, mean, and one standard deviation below the mean. First, we fix the school mean SES (mean = 0, s.d. = 0.41 as in Table 1) at –0.41, 0, and 0.41, and then see the relationship between Math achievement and student's SES (Figure 1).

Insert Figure 1 About Here

As can be seen in (22) and (23), once $\overline{SES}_{.j}$ is fixed, the relationship is linear. The box lines represent schools whose school mean SES is one standard deviation above the mean, i.e., $\overline{SES}_{.j} = 0.41$ (denoted as Catholic+ and public+ in the legend); the triangle lines represent schools whose school mean SES is the grand mean, i.e., $\overline{SES}_{.j} = 0.00$ (denoted as Catholic and public in the legend); and the diamond-shape lines represent schools whose school mean SES is one standard deviation below the mean, i.e., $\overline{SES}_{.j} = -0.41$ (denoted as Catholic - and public - in the legend). The graph shows that from low to middle/high student's SES, going to Catholic school produces higher math achievement, but the gap gets smaller and smaller as the student's SES increases. After passing the certain student's SES, we expect that going to public schools will be a better choice. The value of the student's SES at which public schools catch up to Catholic schools shifts to the right as the school SES increases, i.e., for $\overline{SES}_{.j} = -0.41$, $SES_{ij} = 0.34$, for $\overline{SES}_{.j} = 0$, $SES_{ij} = 0.75$, and for $\overline{SES}_{.j} = 0.41$, $SES_{ij} = 1.16$. This means that the higher the SES of the school, the more students receive the benefits of achieving higher mathematics scores by attending Catholic schools[3].

Though the above description informs us a pattern on how mathematics achievement differ in Catholic and public schools as a function of student's SES given a certain school SES, we are not sure whether those differences are in fact statistically significant. To answer this question, we need to use a statistical test that was described in the previous section. The test, Johnson-Neyman type procedure, determines the sets of the pairs of student's SES and school SES that produce the statistically significant and non-significant differences between Catholic and public schools.

In order to perform the test, we first compute the expected difference between Catholic and public schools. For the same student's SES and school mean SES scores, it is obtained from taking the difference between (22) and (23):

$$E(Y_{ij})_C - E(Y_{ij})_P = \gamma_{02} + \gamma_{12}(SES_{ij} - SES_{.j}). \qquad (26)$$

Thus, the expected difference is the linear function of the relative student's SES score, relative to the school mean SES score. The predicted difference is

$$\hat{Y}_{ij,C} - \hat{Y}_{ij,P} = \hat{\gamma}_{02} + \hat{\gamma}_{12}(SES_{ij} - \overline{SES}_{.j}). \qquad (27)$$

Replacing the parameter values by their estimates in Table 2 provides the numerical relationship:

---

[3] Of course, we can take a look at the Equations (24) and (25) by the other way around, i.e., mathematics achievement as a function of school SES by fixing student's SES. As we can see, once the student's SES is fixed, Equations (24) and (25) become a quadratic function of school SES. We, however, here restrict our attention only on the above case, i.e., mathematics achievement as a function of student's SES given a certain school SES, for the purpose of simplifying the presentation of the idea.

$$\hat{Y}_{ij,C} - \hat{Y}_{ij,P} = 1.23 - 1.64(SES_{ij} - \overline{SES}_{.j}). \tag{28}$$

Thus, the discrepancy decreases as the relative SES increases.

Let the relative student's SES be *x*, i.e., $x = p - q$, where *p* is the student's SES and *q* is the school mean SES. Then, the hypothesis that $E(Y_{ij})_C - E(Y_{ij})_P = 0$ can be cast into the form of linear hypothesis as:

$$\text{H}_0\text{: } K^T\gamma = \mathbf{0} \tag{29}$$

where $K^T = (0,0,1,0,0,x)$ and $\gamma^T = (\gamma_{00,}\gamma_{01},\gamma_{02},\gamma_{10},\gamma_{11},\gamma_{12})$.

Since K^T has one row, the degrees of freedom for the reference chi-square distribution is one. Thus, the individual SES region of significance is given by solving the following inequality with respect to *x:*

$$(K^T\hat{\gamma})^T\hat{V}_K^{-1}(K^T\hat{\gamma}) - \chi^2_{df=1,\alpha} \geq 0 \tag{30}$$

where $\hat{V}_K$ is the estimate of V_K in Equation (11). If we choose 0.05 level, the critical value is 3.84($\chi^2_{df=1,\alpha=0.05}$ = 3.84), we have

$$(K^T\hat{\gamma})^T\hat{V}_K^{-1}(K^T\hat{\gamma}) - 3.84 \geq 0 \tag{31}$$

where $K^T\hat{\gamma} = 1.23 - 1.64x$. Since $\hat{V}_\gamma$ is a six by six symmetric matrix and if we represent $\hat{V}_\gamma$ by elements, i.e.,

$$\hat{V}_\gamma = \begin{pmatrix} \hat{var}(\hat{\gamma}_{00}) & \hat{cov}(\hat{\gamma}_{00},\hat{\gamma}_{01}) & \hat{cov}(\hat{\gamma}_{00},\hat{\gamma}_{02}) & \hat{cov}(\hat{\gamma}_{00},\hat{\gamma}_{10}) & \hat{cov}(\hat{\gamma}_{00},\hat{\gamma}_{11}) & \hat{cov}(\hat{\gamma}_{00},\hat{\gamma}_{12}) \\ \hat{cov}(\hat{\gamma}_{01},\hat{\gamma}_{00}) & \hat{var}(\hat{\gamma}_{01}) & \hat{cov}(\hat{\gamma}_{01},\hat{\gamma}_{02}) & \hat{cov}(\hat{\gamma}_{01},\hat{\gamma}_{10}) & \hat{cov}(\hat{\gamma}_{01},\hat{\gamma}_{11}) & \hat{cov}(\hat{\gamma}_{01},\hat{\gamma}_{12}) \\ \hat{cov}(\hat{\gamma}_{02},\hat{\gamma}_{00}) & \hat{cov}(\hat{\gamma}_{02},\hat{\gamma}_{01}) & \hat{var}(\hat{\gamma}_{02}) & \hat{cov}(\hat{\gamma}_{02},\hat{\gamma}_{10}) & \hat{cov}(\hat{\gamma}_{02},\hat{\gamma}_{11}) & \hat{cov}(\hat{\gamma}_{02},\hat{\gamma}_{12}) \\ \hat{cov}(\hat{\gamma}_{10},\hat{\gamma}_{00}) & \hat{cov}(\hat{\gamma}_{10},\hat{\gamma}_{01}) & \hat{cov}(\hat{\gamma}_{10},\hat{\gamma}_{02}) & \hat{var}(\hat{\gamma}_{10}) & \hat{cov}(\hat{\gamma}_{10},\hat{\gamma}_{11}) & \hat{cov}(\hat{\gamma}_{10},\hat{\gamma}_{12}) \\ \hat{cov}(\hat{\gamma}_{11},\hat{\gamma}_{00}) & \hat{cov}(\hat{\gamma}_{11},\hat{\gamma}_{01}) & \hat{cov}(\hat{\gamma}_{11},\hat{\gamma}_{02}) & \hat{cov}(\hat{\gamma}_{11},\hat{\gamma}_{10}) & \hat{var}(\hat{\gamma}_{11}) & \hat{cov}(\hat{\gamma}_{11},\hat{\gamma}_{12}) \\ \hat{cov}(\hat{\gamma}_{12},\hat{\gamma}_{00}) & \hat{cov}(\hat{\gamma}_{12},\hat{\gamma}_{01}) & \hat{cov}(\hat{\gamma}_{12},\hat{\gamma}_{02}) & \hat{cov}(\hat{\gamma}_{12},\hat{\gamma}_{10}) & \hat{cov}(\hat{\gamma}_{12},\hat{\gamma}_{11}) & \hat{var}(\hat{\gamma}_{12}) \end{pmatrix}, \tag{32}$$

then $\hat{V}_K = K^T \hat{V}_\gamma K = V\hat{a}r(\hat{\gamma}_{02}) + 2x\,c\hat{o}v(\hat{\gamma}_{02}, \hat{\gamma}_{12}) + x^2 V\hat{a}r(\hat{\gamma}_{12})$. The actual values for $\hat{V}_\gamma$ obtained from HLM is:

$$\hat{V}_\gamma = \begin{pmatrix} 3.95*10^{-2} & 1.80*10^{-2} & -4.24*10^{-2} & 2.30*10^{-3} & 1.05*10^{-3} & -2.47*10^{-3} \\ 1.80*10^{-2} & 1.36*10^{-1} & -4.02*10^{-2} & 1.06*10^{-3} & 8.15*10^{-3} & -2.41*10^{-3} \\ -4.24*10^{-2} & -4.02*10^{-2} & 9.38*10^{-2} & -2.47*10^{-3} & -2.40*10^{-3} & 5.63*10^{-3} \\ 2.30*10^{-3} & 1.06*10^{-3} & -2.47*10^{-3} & 2.47*10^{-2} & 1.33*10^{-2} & -2.65*10^{-2} \\ 1.05*10^{-3} & 8.15*10^{-3} & -2.40*10^{-3} & 1.33*10^{-2} & 9.15*10^{-2} & -2.58*10^{-2} \\ -2.47*10^{-3} & -2.41*10^{-3} & 5.63*10^{-3} & -2.65*10^{-2} & -2.58*10^{-2} & 5.90*10^{-2} \end{pmatrix} \quad (33)$$

The contrast matrix K^T involves an unknown value x and thus we use Mathematica to do the symbolic computation.

Step 1: Obtain the estimate of the variance-covariance matrix for $\hat{\gamma}$ when you run HLM by using the keyword "**Print variance-covariance**". This command produces a text file 'gamvc.dat' that contains the variance-covariance matrix of the fixed effects. The matrix (33) is the one we obtained.

Step 2: *Define the estimate of the variance-covariance matrix for* $\hat{\gamma}$, $V\hat{a}r(\hat{\gamma})$: In Mathematika, scientific notation is represented as 10^_. For example, 0.039488440 is 3.9488440 10^_002.

```
Vg = {{3.9488440 10^-002, 1.7959450 10^-002, -4.2425393 10^-002,
   2.2987740 10^-003, 1.0537420 10^-003, -2.4745921 10^-003},
  {1.7959450 10^-002, 1.3628002 10^-001, -4.0245690 10^-002,
   1.0585086 10^-003, 8.1498270 10^-003, -2.4135827 10^-003},
  {-4.2425393 10^-002, -4.0245690 10^-002, 9.3802256 10^-002,
   -2.4743120 10^-003, -2.4032180 10^-003, 5.6273247 10^-003},
  {2.2987740 10^-003, 1.0585086 10^-003, -2.4743120 10^-003,
   2.4686438 10^-002, 1.3342234 10^-002, -2.6503914 10^-002},
  {1.0537420 10^-003, 8.1498270 10^-003, -2.4032180 10^-003,
   1.3342234 10^-002, 9.1546401 10^-002, -2.5814267 10^-002},
  {-2.4745921 10^-003, -2.4135827 10^-003, 5.6273247 10^-003,
   -2.6503914 10^-002, -2.5814267 10^-002, 5.9003012 10^-002}}
```

Step 3: *Define the vector* $\hat{\gamma}$, the estimate of the fixed effects.

```
g = {{12.0950064}, {5.3330565}, {1.2263840}, {2.9377875}, {1.0344270}, {-1.6409540}}
```

Note that in order to define the column vector, we enclose each elemets by a blacket { }.

Step 4: *Define the contrast matrix*:

```
K' = {{0, 0, 1, 0, 0, x}}
```

Step 5: Compute the variance-covariance matrix of $K^T\hat{\gamma}$.

```
VK = K'.Vg.Transpose[K']
```

The output provided by Mathematica for this operation is given below:

$\{\{0.0938023 + 0.00562732\,x + (0.00562732 + 0.059003\,x)\,x\}\}$.

Mathematica can simplify this result by using the command **Simplify.**

```
Simplify[VK]
```

$\{\{0.0938023 + 0.0112546\,x + 0.059003\,x^2\}\}$

Step 6: compute the statistics *H*.

```
H = Transpose[g].Transpose[K'].Inverse[VK].K'.g
```

$$\left\{\left\{\frac{1.22638\ (1.22638-1.64095\,x)}{0.0938023+0.00562732\,x+(0.00562732+0.059003\,x)\,x}-\frac{1.64095\ (1.22638-1.64095\,x)\,x}{0.0938023+0.00562732\,x+(0.00562732+0.059003\,x)\,x}\right\}\right\}.$$

Again, a more concise representation of this equation can be obtained by asking for a simplification:

```
Simplify[H]
```

$$\left\{\left\{\frac{2.69273\ (-0.74736+x)\ (-0.74736+x)}{0.0938023+0.0112546\,x+0.059003\,x^2}\right\}\right\}.$$

As can be seen from the above expression, Mathematica represents $H=(K^T\hat{\gamma})^T\hat{V}_c^{-1}(K^T\hat{\gamma})$ as the ratio of a quadratic polynomial both in the numerator and in the denominator.

Step 7: Solve the equation $H-3.84=0$ for the unknown *x*.

```
Solve[H - 3.84 == 0, x]
```

and the results provided by Mathematica are as follows:

```
{{x → 0.359527}, {x → 1.29004}}
```

for which the two real roots 0.359527 and 1.29004 represent the solution of interest.

Step 8. Plot the function (H – 3.84) over the range of $-4 \le x \le 4$:

```
Plot[H - 3.84, {x, -4, 3}, AxesLabel → {"R. SES", "H-3.84"}]
```

where ‘R. SES’ represents the student’s SES relative to the school mean. The plot is given in Fig. 2.

Insert Figure 2 About Here

The axes labeled as $H - 3.84$ represents the value of $\hat{\gamma}^T K \hat{V}_K^{-1} K^T \hat{\gamma}$ - 3.84 for various values of x, the relative student's SES.

The above results analyzed via Mathematica package indicate that significance is attained when the groups are equated to a value of the student's relative SES set to 0.36 or smallest and 1.29 or greater; i.e., the point at which the function crosses the horizontal axis. Significant differences at $\alpha < 0.05$ will be found for all values of x for which the function plotted is above the horizontal axis. With 95 % confidence, we conclude that when the student's relative SES is less than 0.36 the student will be expected to do better in Catholic schools than public schools. On the other hand, if the student's relative SES is greater than 1.29, the student will do better in public schools than in Catholic schools. If the student's relative SES is between 0.36 and 1.29, there is no statistically significant difference whether student chooses Catholic high school or public high school.

Another way of examining the student's relative SES range that produces either significant or nonsignificant differences is to directly examine the expected difference between Catholic and public schools, represented by Equation (26) or $K^T\gamma$ where $K^T = (0,0,1,0,0,x)$ for x be the student's relative SES and $\gamma^T = (\gamma_{00,}\gamma_{01}, \gamma_{02}, \gamma_{10}, \gamma_{11}, \gamma_{12})$. Since a contrast K^T in this case has one row, a 95 % confidence interval on $K^T\gamma$ can be consructed by a formula

$$K^T \hat{\gamma} \pm \sqrt{3.84 \hat{V}_K} \tag{34}$$

where 3.84 is critical value at the upper 0.05 cutoff of chi-square variate of one degrees of freedom, $K^T \hat{\gamma}$ is the predicted difference as in Equation (27) or Equation (28), and $\hat{V}_K$ is the estimate of the $Var(K^T \hat{\gamma})$, which is represented in Equation (11).

The predicted difference and its confidence interval can be best described by the graph.

Insert Figure 3 About Here

The bold straight line in Figure 3 represents the predicted difference, expressed by Equation (28), or symbolically, by $K^T \hat{\gamma}$. The upper curve represents the upper bound of the 95 % confidence interval and the lower curve represents the corresponding lower bound[4]. From the figure, we find that the expected mathematics achievement of students in Catholic schools is higher in the region of $(-\infty, 0.75)$ in relative student's SES than that in public schools, where 0.75 is the value of the relative student's SES at which the bold line crosses the horizontal axis and this value was the same value that was observed in Figure 1, and then the mean achievement of public schools gets higher in the region of $(0.75, \infty)$ than that of catholic schools. To see if the difference is statistically significant or not, we focus on which values the lower bound line and the upper bound line cross the abscissa. They are 0.36 for the line of lower bound and 1.29 for the upper bound line.

[4] Note that, strictly speaking, this confidence interval is not simultaneous one. See the comments regarding this in the Summary and Discussion section.

Thus, we conclude that the expected mathematics achievement of students in Catholic schools are statistically significantly higher in the region of ($-\infty$, 0.36) in relative student's SES than that in public schools. On the other hand, the expected mathematics achievement of students in public schools are statistically significantly higher in the region of (1.29, $+\infty$) in relative student's SES than that in Catholic schools.

Out of all 7185 students in our sample, 4993 students (69.5 %) have their relative SES lower than 0.36, 2064 students (28.7 %) have their SES between 0.36 and 1.29, and 128 students (1.8 %) has the relative SES higher than 1.29. Thus, it can be said that, on average, about 70 % of the students will gain better math scores if they choose to go to a Catholic school rather than a public school when two sectors have the same school SES.

If we classify those three groups of students by the sector (Table 3), then we find that 2469 for Catholic schools and 2524 for public schools out of 4993 students whose relative SES is lower than 0.36. Thus, it can be said that 69.6 % (2469/3543) of the students in Catholic school made the right decision, and 69.3 % (2524/3642) of the students in public schools could have obtained the better math scores if they went to Catholic schools instead of public schools. For high relative SES students, 93 students (2.6% of the all public school students) in public schools had their relative students SES (RSES) greater than or equal to 1.29, and thus they benefited by having gone to public schools; 35 students (1.0% of the all Catholic school students) in Catholic schools were disadvantaged by having gone to Catholic schools. In overall, 2562 (35.7%) students made a right choice in terms of the choice of the sector, and 2559 (35.6%) students could have gained better math achievement if they have chosen the other sector of the high schools. For the rest of the 2064 (28.7%) students, the sector did not create the statistically significant difference. As a conclusion from the above analysis, we might make the following comments if we were asked an advice from the students and the parents who were considering the sector of the schools in order to achieve the highest possible mathematics scores: Choose a school whose school SES is as high as possible. If you find the schools that have the same school SES in both Catholic and public sectors, then consider your family SES. Choose a Catholic school if your family SES is lower than the school SES or is higher up to 0.36 than the school SES. On the other hand, choose a public school if your family SES is extremely higher than the school SES, specifically 1.29. If your family SES relative to the school SES is in between 0.36 and 1.29, then the sector does not matter. For more than two third of the students (69.5 %), it would be more advantageous to choose a Catholic sector.

Though it's possible to interpret the results directly by the relative student's SES as was done in the above, translating the relative student's SES back into the original scales, i.e., student's SES and school mean SES, provides an extra insight. Since $x = p - q$ and the non-significance region is represented by the inequality, $0.36 \le x \le 1.29$, the information is mapped into the two dimensional space, shown in Figure 4.

Insert Figure 4 About Here

The large rectangle delineates the data points in the sample and the two parallel lines, $q = p - 0.36$ and $q = p - 1.29$ divides the plane into three regions. The upper left area in dark shadow represents the region that Catholic schools do better than public schools significantly. The area that lies between the parallel lines indicates the region that there is no statistically significant difference between two sectors in this region. The right below area is the region that public schools will produce better results than Catholic schools. We can use this figure to answer some interesting questions. For example, let us consider a student with certain SES and who is looking for which sector of high school he/she should attend in order to achieve better in mathematics. Suppose his/her SES is 0.5, which is a little lower than the one standard deviation high above the mean (See the vertical line at SES = 0.5 in Fig. 4). The answer depends on what level of school SES is available. The student does better in public school if he/she goes to a school whose school SES is lower than –0.79; there would be no significant difference if the school SES is between –0.79 and 0.14; and the student does better in Catholic school if he/she goes to a school whose school SES is higher than 0.14. Of course, to maximize the math achievement, higher the school SES, the better his/her math score will be[5].

We can use Figure 4 by the other way around. That is, suppose a situation that an local administrator of a school district whose high schools have the same school SES ponders what kind of students can benefit by going to either Catholic or public sector. Assume, for example, the school SES is fixed at the grand mean, 0. Then up to 0.36 of student's SES, those students will do better if they are sent to Catholic schools. There will be no statistical difference if the student's SES is between 0.36 and 1.29. And if student's SES is higher than 1.29, then he/she will make better performance if he/she goes to public school.

[5] Since the predicted math achievement is the quadratic function with respect to school SES (see Equations (24) and (25) with the coefficient for the quadratic term be negative), there is the school SES that

Now, let us consider a slightly different question. Suppose we want to know about the range of student's SES that produces the significant difference between Catholic and public sector in terms of mathematics achievement if the two students who have the same individual *SES* are assigned to a typical Catholic school and a typical public school. The "typical" indicates that the school SES is the mean SES of all the schools in the same sector. Let the individual SES be denoted *x*, the mean school SES of all the Catholic schools be s_C, and the mean school SES of all the public schools be s_P. In our data, those values are $s_C = 0.166$ and $s_P = -0.130$ for each. The expected values for those students in each typical school are obtained from Equations (22) and (23) by setting $\overline{SES}_{.j} = s_C$ for Catholic schools and $\overline{SES}_{.j} = s_P$ for public schools:

$$\begin{aligned} E(Y_{ij})_{Typ.\,Cath.} &= \{\gamma_{00} + \gamma_{02} + \gamma_{01} s_C\} + \{\gamma_{10} + \gamma_{12} + \gamma_{11} s_C\}(x - s_C) \\ &= \{\gamma_{00} + (\gamma_{01} - \gamma_{10}) s_C + \gamma_{02} - \gamma_{11} s_C^2 - \gamma_{12} s_C\} + (\gamma_{10} + \gamma_{11} s_C + \gamma_{12}) x. \end{aligned} \tag{34}$$

$$\begin{aligned} E(Y_{ij})_{Typ.\,Pub.} &= \{\gamma_{00} + \gamma_{01} s_P\} + \{\gamma_{10} + \gamma_{11} s_P\}(x - s_P) \\ &= \{\gamma_{00} + (\gamma_{01} - \gamma_{10}) s_P - \gamma_{11} s_P^2\} + (\gamma_{10} + \gamma_{11} s_P) x \end{aligned} \tag{35}$$

The predicted values for each sector are obtained by replacing the parameter values by their estimates and $s_C = 0.1661$ and $s_P = -0.1295$:

$$\begin{aligned} \hat{Y}_{ij,Typ.\,Cath.} &= \{\hat{\gamma}_{00} + (\hat{\gamma}_{01} - \hat{\gamma}_{10}) s_C + \hat{\gamma}_{02} - \hat{\gamma}_{11} s_C^2 - \hat{\gamma}_{12} s_C\} + (\hat{\gamma}_{10} + \hat{\gamma}_{11} s_C + \hat{\gamma}_{12}) x \\ &= 13.963 + 1.649x \end{aligned} \tag{36}$$

and

$$\begin{aligned} \hat{Y}_{ij\,Typ.\,Pub.} &= \{\hat{\gamma}_{00} + (\hat{\gamma}_{01} - \hat{\gamma}_{10}) s_P - \hat{\gamma}_{11} s_P^2\} + (\hat{\gamma}_{10} + \hat{\gamma}_{11} s_P) x \\ &= 11.768 + 2.804x \end{aligned} \tag{37}$$

---

maximizes the student's math achievement. For the student whose SES is 0.5, it is 2.20. But there is no school whose school SES is 2.20 in our data set.

The relationships between the outcome (Math achievement score) and the individual SES for each sector are depicted in Figure 5.

Insert Figure 5 About Here

From the figure, we see that at about SES = 1.9, the two average lines intersect each other. The exact value is 1.900, obtained by equating Equations (36) and (37), and this intersection point sifted to the right, compared to all of the three student's SES values of the intersection points (0.34, 0.75, and 1.16) in Figure 1. In the region where SES < 1.9, persons in the Catholic schools have higher math achievement on average and in the region where SES > 1.9, persons in public schools have higher math achievement than the person in Catholic school given the same individual SES. Since there are only two students who have SES > 1.9 out of 7185 (0.0003 %), both of whom happened to be in public schools, it is clear that, in general, it is more advantageous to go to typical Catholic schools than typical public schools in order to obtain higher math achievement.

Now the question is, "In which region of SES two group means are statistically significantly different and which group is higher in that range and in which region of SES there is no difference statistically?" In order to answer the question, we use the Johnson-Neyman type technique again. As before, we first obtain the expected difference from Equation (35) and Equation (36), which is

$$E(Y_{ij})_{Typ.Cath.} - E(Y_{ij})_{Typ.Pub.} = \{(\gamma_{01} - \gamma_{10})(s_C - s_P) + \gamma_{02} - \gamma_{11}(s_C^2 - s_P^2) - \gamma_{12}s_C\} + \{\gamma_{11}(s_C - s_P) + \gamma_{12}\}x \tag{39}$$

Thus the hypothesis that $E(Y_{ij})_{Typ.\,Cath.} - E(Y_{ij})_{Typ.\,Pub.} = 0$ can be expressed in the form of Equation (9), i.e., a linear combination of the fixed effects paramters as:

$$\mathrm{H_0}\text{: } K^T\gamma = \mathbf{0}$$

where $K^T = [0, s_C - s_P, 1, -(s_C - s_P), (s_C - s_P)(x - s_C - s_P), x - s_C]$ and $\gamma^T = (\gamma_{00}, \gamma_{01}, \gamma_{02}, \gamma_{10}, \gamma_{11}, \gamma_{12})$. Since K^T has one row, the degrees of freedom for the reference chi-square distribution is one. Thus, the individual SES region of significance is given by solving the following inequality with respect to *x:*

$$(K^T\hat{\gamma})^T \hat{V}_K^{-1}(K^T\hat{\gamma}) - \chi^2_{df=1,\alpha} \geq 0 \tag{40}$$

If we choose 0.05 level, the critical value is 3.84($\chi^2_{df=1,\alpha=0.05}$ = 3.84), and in our data, since $s_C = 0.1661$ and $s_P = -0.1295$, we have

$(K^T\hat{\gamma})^T \hat{V}_K^{-1}(K^T\hat{\gamma}) - 3.84 \geq 0$ where

$$K^T = [0, 0.2956, 1, -0.2956, 0.2956(x - 0.0366), x - 0.1661]. \tag{41}$$

The contrast matrix K^T involves an unknown value x and thus we use Mathematica to do the symbolic computation. Following the same steps as in the previous computation, we reach the solution for Equation (40):

$$x \leq 1.07738 \text{ or } x \geq 2.63906.$$

The results indicate that the significant region shifts more to the right, higher student SES, compared to the case when two sectors have the same school SES. This happens because the mean school SES is higher in Catholic schools than public schools, i.e., $s_C = 0.166$ and $s_P = -0.130$, and the estimated effects of school SES ($\hat{\gamma}_{01}$ and $\hat{\gamma}_{11}$) are positive both for the intercept ($\hat{\beta}_{0j}$) and the student's SES slope ($\hat{\beta}_{1j}$).

The significance is attained when the groups are equated to a value of the individual SES set to 1.08 or smallest and 2.64 or greater. With 95 % confidence, we can say that when the student's SES is less than 1.08 the student will be expected to do better in Catholic schools than public schools. On the other hand, if the student's SES is greater than 2.64, the student will do better in public schools than in Catholic schools. If the student's SES is between 1.077 and 2.639, whose interval includes 1.90, which is the intersection point (Fig. 5), whether student chooses Catholic high school or public high school makes no statistically significant difference. Out of all 7185 students, 6572 students (91.5 %) have their SES lower than 1.077, 612 students (8.5 %) have their SES between 1.077 and 2.639, and only one student (0.014 %) has the SES higher than 2.639. Thus, it can be said that, based on student's SES in the sample, more than 90 % of the students will gain significantly better mathematics scores in typical Catholic schools than in typical public schools.

Example 2. Early Child Cognitive Development

In a longitudinal growth study, there are repeated measures for the same subjects. In multilevel modeling framework, we conceptualize the problem in a way that each individual has his/her own growth trajectory. Suppose at level-1, a polynomial is formulated to model each individual growth on an outcome measure over time. The level-1 coefficients are supposed to vary from person to person. At level-2, we frequently try to explain those variability by some individual characteristics such as gender, race, and so forth. Those individual characteristics that have qualitative nature create the groups whose members share the same characteristics. As a result, each group has mean growth trajectory, which differ in terms of the intercepts and age-related slopes. In this context, a question such as "In what age range, two groups are statistically different?" may be of the interest. Let us consider an example.

The Greensboro early child cognitive development study (Chirstian, Bachman, & Morrison, 2001) aims to study developmental trajectory and schooling effects on cognitive ability on young children whose grades ranged from kindergarten to third grade. In order to disentangle age-related and schooling-related influences on children's cognitive growth, the study employs "cut-off" methods, a natural experiment to take advantage of school cutoff. Most North American school district designate an arbitrary "cutoff" date such that children whose birthdate precedes the date will be allowed to entry to kindergarten or first-grade, whereas children whose birthdate just misses the date will be denied entry. The former group of children who entered kindergarten and whose age are younger than peers are called "Young" kindergartner and the latter group of children are called "Old" kindergartner because their age are relatively older than peers. Between the two groups above, there is another group called "Middle" kindergartner whose chronological age was between "Young" and "Old". Other namings such as "Young" 1st grader, "Middle" 1st grader, "Old" 1st grader "Young" 2nd grader, "Middle" 2nd grader, "Old" 2nd grader, etc. were used in the same way. Thus, the design of the study allows us to investigate the net effects of schooling by comparing, for example, "Young" 1st grader and "Old" Kindergartner, because "Young" 1st grader already received one year schooling and "Old" Kindergartner didn't receive it though they have almost the same chronological age. In terms of age-related growth, we can use repeated measures of children who are in the same group. For our analysis, there are 493 children and 173 children (35.1 %) are from "Old" group, 187 (37.9 %) are from "Middle" group, and 133 (27.0 %) are from "Young" group. In terms of race/ethnicity, 254 of them were Caucasians and 239 of them are African Americans. Thus, 48.5 % of the children are African American and it was oversampled so that we can obtain the precise estimates for this group. Also, there are three cohorts depending on the year of entry to the study, i.e., 1990-91, 1991-92, and 1992-93 cohorts. Note that "Middle" Kindergartner was employed to the study from the 1992-93 study, no observations were available to the 1990-91 and 1991-92 cohorts[6]. There were five waves of observations and the first observation was made at fall of kindergarten; the second observation was at spring of kindergarten; the third was at spring of 1st grade; the fourth was at spring of 2nd grade; and the fifth was at spring of 3rd grade. At the time of 1998, this study was continuing, so the observations for the spring of 3rd grade were not made yet for the 1992-93 cohorts. Also, there are missing observations because of attrition; at the spring of 3rd grade, 39 % had complete observations for 1990-91 and 1991-92 cohorts. The descriptive statistics for the key variables are provided in Table 4.

Insert Table 4 About Here

In order to study age-related cognitive growth, a two-level hierarchical linear model was employed. The level-1 model was for child j at time:

[6] Since no significant cohort and cohort-grade interaction effects were found, cohort indicators are

L-1:

$$(Math)_{ij} = \beta_{0j} + \beta_{1j}(Grade)_{ij} + \beta_{2j}(Grade)_{ij}^2 + \beta_{3j}(DFall)_{ij} + \varepsilon_{ij}, \qquad (42)$$

where $\varepsilon_{ij} \overset{iid}{\sim} N(0,\sigma^2)$ and $Math_{ij}$ is the mathematics score for child *j* at time *i*; $Grade_{ij}$ is the child *j*'s grade at time *i* and it is coded as –1.5 for fall of kindergarten, -1.0 for spring of kindergarten, 0 for spring of 1st grade, 1.0 for spring of 2nd grade, and 2.0 for spring of 3rd grade; $Grade_{ij}^2$ is the square of $Grade_{ij}$; $DFall_{ij}$ is the indicator of fall term measurement and thus it is equal to 1 if the math score was measured at fall and 0 if not; and ε_{ij} is the random within-child error. Note that the reason that the indicator of fall measurement was included is that we wanted to capture yearly growth trajectory, which includes summer term that doesn't have school, by adjusting the fall of kindergarten measurement.

The level-2 model represents the variability of the level-1 random regression coefficients.

L-2:

$$\beta_{0j} = \gamma_{00} + \gamma_{01}(DBlack)_j + \gamma_{02}(IQ_j - \overline{IQ}_.) + \gamma_{03}(LIT_j - \overline{LIT}_.) + \gamma_{04}(DO)_j + \gamma_{05}(DM)_j + \gamma_{06}(MOMED_j - \overline{MOMED}_.) + u_{0j}$$

$$\beta_{1j} = \gamma_{10} + \gamma_{11}(DBlack)_j + \gamma_{12}(IQ_j - \overline{IQ}_.) + \gamma_{13}(DO)_j + \gamma_{14}(DM)_j + u_{1j} \qquad (43)$$

$$\beta_{2j} = \gamma_{20} + \gamma_{21}(DBlack)_j + \gamma_{22}(IQ_j - \overline{IQ}_.) + \gamma_{23}(DO)_j + \gamma_{24}(DM)_j + u_{2j}$$

$$\beta_{3j} = \gamma_{30}$$

where

completely dropped from the subsequent analyses.

$$\begin{pmatrix} u_{0j} \\ u_{1j} \\ u_{2j} \end{pmatrix} \sim N(\begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} \tau_{00} & \tau_{01} & \tau_{02} \\ \tau_{10} & \tau_{11} & \tau_{12} \\ \tau_{20} & \tau_{21} & \tau_{22} \end{pmatrix});$$

DBlack is the indicator for African American and it is 1 if the child is African American and 0 if not; *IQ* is the child IQ measured at the fall of the kindergarden and it is centered around the grand mean; *LIT* is the child's home literacy environment measure and it is centered around the grand mean; *DO* is the indicator for the children in the "Old" group; *DM* is the indicator for the children in the "Middle" group ; and *MOMED* is the mother's years of education and it is centered around the grand mean; γ_{00} represents the expected math achievement score at the spring of 1st grade for the white child in the "Young" group whose IQ, home literacy and mother's years of education are the average; γ_{10} represents the expected slope at the spring of 1st grade for the white child in the "Young" group whose IQ, home literacy and mother's years of education are at their averages; γ_{20} represents the expected rate of acceleration for the white child in the "Young" group whose IQ, home literacy and mother's years education are at their averages; and γ_{30} represents the mean discrepancy of the fall measurement from the general growth trajectory.

In this model, since there are two groups for race/ethnicity (white and black represented by DBLACK variable) and three levels of chronological groups ("Young", "Middle", and "Old", represented by two dummy variables "DO" and "DM"), we have six trajectories which represent the average growth of mathematics ability from fall of kindergarten to spring of 3rd grade.

The results for the model in Equations (42) and (43) are in the Table 5.

Insert Table 5 About Here

From Table 5, the average growth trajectories for six groups are all different even after controlling for IQ, home literacy environment, and mother's years of education.

Now suppose we are interested in the growth trajectories of white Young children and black Old children. The expected value for the children in those groups at the average IQ, the average home literacy environment, and the average mother's years education are for black Old children:

$$\begin{aligned} E(Y_{ij})_{BO} = {} & (\gamma_{00} + \gamma_{01} + \gamma_{04}) + (\gamma_{10} + \gamma_{11} + \gamma_{13})(Grade)_{ij} + (\gamma_{20} + \gamma_{21} + \gamma_{23})(Grade)^2_{ij} \\ & + \gamma_{30}(DFall)_{ij} \end{aligned} \quad (44)$$

and for white Young children:

$$E(Y_{ij})_{WY} = (\gamma_{00}) + (\gamma_{10})(Grade)_{ij} + (\gamma_{20})(Grade)^2_{ij} + \gamma_{30}(DFall)_{ij} \quad (45)$$

Therefore, at the same grade, the expected mean difference is

$$E(Y_{ij})_{BO} - E(Y_{ij})_{WY} = (\gamma_{01} + \gamma_{04}) + (\gamma_{11} + \gamma_{13})(Grade)_{ij} + (\gamma_{21} + \gamma_{23})(Grade)^2_{ij}. \quad (46)$$

The predicted values can be computed by the following equations for the two groups.

$$\begin{aligned} \hat{Y}_{ij\,BO} = {} & (\hat{\gamma}_{00} + \hat{\gamma}_{01} + \hat{\gamma}_{04}) + (\hat{\gamma}_{10} + \hat{\gamma}_{11} + \hat{\gamma}_{13})(Grade)_{ij} + (\hat{\gamma}_{20} + \hat{\gamma}_{21} + \hat{\gamma}_{23})(Grade)^2_{ij} \\ & + \hat{\gamma}_{30}(DFall)_{ij} \\ \hat{Y}_{ij\,WY} = {} & \hat{\gamma}_{00} + \hat{\gamma}_{10}(Grade)_{ij} + \hat{\gamma}_{20}(Grade)^2_{ij} + \hat{\gamma}_{30}(DFall)_{ij} \end{aligned} \quad (47)$$

Replacing the values of the parameters by their estimated values provided in Table 4, we obtain the predicted trajectories for the two groups. Those are:

$$\hat{Y}_{ij,BO} = 26.463 + 8.284 Grade_{ij} + 0.142 Grade^2_{ij} - 1.380 DFall_{ij}$$

$$\hat{Y}_{ij,WY} = 22.179 + 10.115Grade_{ij} + 2.004Grade_{ij}^2 - 1.380DFall_{ij}. \qquad (48)$$

In Figure 6, the two predicted curves are depicted as the function of the grade.

Insert Figure 6 About Here

As can be seen from the figure, the two curves are crossing just before the spring of the 2nd grade. Before that, the black children in the "old" group had higher math scores than white children in the "young" group, but after the spring of 2nd grade, white "young" children surpassed the black "old" children. Specifically, at fall of kindergarten, black old group is 2.84 higher than white young group and this difference is statistically significant (see Table 6). At the spring of kindergarten, the difference increases to 4.25, and it stays at the almost same value (4.28) until the spring of 1st grade. From this point, the superiority of black old group starts diminishing. At the spring of the 2nd grade, the difference goes down to 0.52, which is not statistically significant anymore. Finally, at the spring of 3rd grade, white young children do significantly better than black old children in mathematics by 6.83, which is the largest margin among the five time points.

Insert Table 6 About Here

Though we can test whether there is a statistically significant difference at each time point, one at a time, by specifying the appropriate contrast as seen in Table 4, this procedure is cumbersome, especially when the data have many time points. If we use the

Johnson-Neyman type test developed in this article with the symbolic computation capability of Mathematica, we can do all the tests at once. That is, to test the hypothesis that $E(Y_{ij})_{BOK} - E(Y_{ij})_{WYK} = 0$ at Grade t can be expressed in the form of Equation (9), i.e., a linear combination of the fixed effects paramters as:

$$\mathrm{H}_0\text{: } K^T \gamma = \mathbf{0} \tag{49}$$

where $K^T = [0,1,0,0,1,0,0,0,t,0,t,0,0,t^2,0,t^2,0,0]$ and

$\gamma^T = (\gamma_{00}.\gamma_{01},\gamma_{02},\gamma_{03},\gamma_{04},\gamma_{05},\gamma_{06},\gamma_{10},\gamma_{11},\gamma_{12},\gamma_{13},\gamma_{14},\gamma_{20},\gamma_{21},\gamma_{22},\gamma_{23},\gamma_{24},\gamma_{30})$.

Since K^T has one row, the degrees of freedom for the reference chi-square distribution is one. Thus, the Grade region of significance is given by solving the following inequality with respect to *t:*

$$\hat{\gamma}^T K \hat{V}_K^{-1} K^T \hat{\gamma} - \chi^2_{df=1,\alpha} \geq 0.$$

If we choose 0.05 level, the critical value is 3.84($\chi^2_{df=1,\alpha=0.05} = 3.84$), we have

$$\hat{\gamma}^T K \hat{V}_K^{-1} K^T \hat{\gamma} - 3.84 \geq 0.$$

The same steps can be used to decide the region of significance on grade. In Step 6 of the previous example, we computed the $H = \hat{\gamma}^T K \hat{V}_K^{-1} K^T \hat{\gamma}$ statistic. Simplifying the expression of H via Mathematica, we obtain

```
Simplify[H]
```

$$\left\{\left\{\frac{10.4227\,(2.08592+t)\,(2.08592+t)\,(1.21617-2.2056\,t+t^2)}{(2.30517-1.73069\,t+t^2)\,(2.06208+2.71154\,t+t^2)}\right\}\right\}.$$

Mathematica represents H as the ratio of a quadraric polynomial both in the numerator and in the denominator. In step 7, we solve the equation $H - 3.84 = 0$ for the unknown t.

```
Solve[H- 3.84 == 0, t]
```

and the results provided by Mathematica are as follows:

```
{{t→ -3.26576}, {t→ -1.71801}, {t→ 0.519488}, {t→ 1.92323}}
```

for which the four real roots –3.27, -1.72, 0.52 and 1.92 represent the solution of interest. Plotting the function (H – 3.84) over the range of $-5 \le t \le 3$ makes clear on which region of grade are really significant.

```
Plot[H- 3.84, {t, -5, 3}, AxesLabel→ {"Coded Grade", "H-3.84"}]
```

Insert Figure 7 About Here

The axes labeled as $H - 3.84$ represents the value of $\hat{\gamma}^T K \hat{V}_K^{-1} K^T \hat{\gamma}$ - 3.84 for various values of t, the child's Grade.

The results indicate that significance is attained when the groups are equated to a value of the grade set to –3.27 or smallest, between -1.72 and 0.52, and 1.92 or greatest. Translating back to original grade, we identify that at the fall of kindergarten, spring of kindergarten, spring of 1st grade, and spring of 3rd grade the expected mathematics scores of the two groups are statistically different, and only at the spring of 2nd grade, there is no statistical difference on mathematics scores between the two groups.

An interesting observation can be made in Table 6. Comparison the chi-square values of spring of kindergarten and spring of 1st grade, we find that even though the predicted differences are about the same (4.25 vs. 4.28), the chi-squares have the large difference (30.79 vs. 11.60). Also, at the spring of 3rd grade, though the predicted difference is the largest among five time points (-6.83), the chi-square is rather small (4.29) and it marginally achieved the significant (P-value = 0.036). These results were occurred mainly because of the attrition of the samples. See Table 4-a) that the sample size for the combined group of White Young and Black Old changes 119, 117, 89, 57, and 27 as time goes from fall of kindergarten, spring of kindergarten, spring of 1st grade, spring of 2nd grade, and spring of 3rd grade. We can find the sudden drop in sample size from spring of kindergarten to spring of 1st grade, and from spring of 2nd grade to spring of 3rd grade.

Actually, $K^T\gamma$ in (49) represents the expected difference between White Young and Black Old children and $K^T\hat{\gamma}$ in $H = \hat{\gamma}^T K \hat{V}_K^{-1} K^T \hat{\gamma}$ represents the predicted difference or we might say the effect size and $\sqrt{\hat{V}_K}$ represents the estimated standard error for the $K^T\gamma$, the expected difference. Then, since the estimated difference $K^T\hat{\gamma}$ and the estimated variance $\hat{V}_K$ are the functions of GRADE, we can see how these changes with GRADE by plotting via MATHEMATICA.

```
ES = K'.g
{{4.28356 - 1.8307 t - 1.86214 t^2}}
```

where ES, K', and *g* represent the effect size $K^T\hat{\gamma}$, the contrast K', and the estimates of the fixed effects $\hat{\gamma}$. The letter *t* in the equation represents the coded grade. The plot of this function over the range of $-3 \le t \le 3$ gives Figure 8.

Insert Figure 8 About Here

We can see the value at fall of kindergarten by:

```
ES1 = ES /. t→ -1.5
```

```
{{2.8398}}.
```

The symbol "/. ->" is the MATHEMATICA function that provides the value at t = –1.5. Similarly, at other time points, we obtain:

```
ES2 = ES /. t→ -1
```

```
{{4.25212}}.
```

```
ES3 = ES /. t→ 0
```

```
{{4.28356}}.
```

```
ES4 = ES /. t→ 1
```

```
{{0.590728}}.
```

```
ES5 = ES /. t→ 2
```

```
{{-6.82637}}.
```

These agree with the values in Table 6.

Now, since $\hat{V}_K$ is a scalar in this example, we can rewrite H as

$$H = \hat{\gamma}^T K \hat{V}_K^{-1} K^T \hat{\gamma} = \frac{(K^T \hat{\gamma})^2}{\hat{V}_K}.$$

Then we can investigate which part of the equation, i.e., numerator or denominator, dominate the value of H, a chi-square test statistic. The investigation is performed in the following:

The plot of the numerator, $(K^T \hat{\gamma})^2$, the square of the effect size, over the range of $-3 \le t \le 3$ is:

Insert Figure 9 About Here

The plot of the denominator, $\hat{V}_K$, the estimated variance or the square of the estimated standard error is obtained by the following steps.

First, compute the estimated variance.

```
VK = K'.Vg.Transpose[K']
```

where V_g is $\hat{Var}(\hat{\gamma})$ and the numeric values are provided by HLM software. The MATHEMATICA output is:

```
{{1.58143 + 0.446094 t - 0.346415 t^2 +
   t (0.237645 + 0.289965 t + 0.0709167 t^2) + t (0.20845 + 0.294547 t + 0.0922432 t^2) +
   t^2 (-0.175341 + 0.0702023 t + 0.159011 t^2) + t^2 (-0.171074 + 0.0929576 t + 0.173681 t^2)}}.
```

We simplify this function as before.

```
Simplify[VK]
```

```
{{0.332692 (2.30517 - 1.73069 t + t^2) (2.06208 + 2.71154 t + t^2)}}.
```

The plot of $\hat{V}_K$ over the range of $-3 \le t \le 3$ is in Figure 10.

Insert Figure 10 About Here

Summary and Discussion

Johnson-Neyman technique was developed to determine the covariate's region of significance when the parallel slopes assumption does not hold in an ANCOVA context. Since, in the ANOCOVA context, the model can be formulated by a linear model, we can define the Johnson-Neyman problem by an equation that F-statistic must satisfy, and the roots of the equation provide the solution. The practical problem to obtain the roots of the equation was solved by using Mathematica's capability of symbolic computation. When the model is multilevel, unfortunately the F-test, that is exact even for small samples for the linear model, is inappropriate because the within group errors are not identical. Instead, we can use the approximate Wald test which is asymptotically distributed as a chi-square distribution and thus which would be valid when the sample size, especially the number of clusters, is large enough.

In the first example, we illustrated how Johnson-Neyman type technique was useful in organizational research where the key covariate is continuous. This example is a straightforward extension of Johnson-Neyman technique to multilevel modeling settings. Identifying the specific regions of the pairs of individual SES and school SES, which produce the statistically significant difference, provides the more precise information on which sector of schools is advantageous to achieve the higher math scores.

In the second example, the Johnson-Neyman type technique was applied to a growth model where measurements were made at several discrete ages/grades, though those can be theoretically considered as continuous variables. It showed that the technique was useful to determine in which time points which group achieves significantly higher math scores than the other group, all at once. This all-at-once procedure reduces the amount of tasks required to test at each time point. Also, it should be noted that the growth curves were quadratic curves, not linear, could be handled by the exactly same way as the straight-line case just by changing the contrast matrix in the linear hypothesis.

Finally, it should be mentioned that, in this article, simultaneous confidence approach, as provided by Potthoff (1964, 1983) in the linear model scenario, was not considered. The simultaneous confidence approach will be required when we want to hold the probability of making a Type I error at α for all statements simultaneously. Then, either the Bonferroni or the Scheffé's multiple comparison procedure should be considered. In either way, the simultaneous inference will put a larger critical value of the chi-square distribution to protect the test from inflated Type-I error. The detailed study will be left to the future study.

References

Bryk, A. S., & Raudenbush, S. W. (1992). *Hierarchical Linear Models. Applications and Data Analysis Methods*. Newbury Park, CA: Sage.

Christian, K., Bachman, H. J., & Morrison, F. J. (2001). Schooling and Cognitive Development. In R. J. Steinberg & E. L. Grigorenks (Ed.) *Environmental effects on cognitive abilities.* Mahwah, NJ: Lawrence Erlbaum Associates, pp 287-336.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences (2nd ed.).* Hillsdale, New Jersey: Lawrence Erlbaum.

Coleman, J. S., Hoffer, T., Kilgore, S. B. (1982). *High school achievement: Public, Catholic and other schools compared.* New York: Basic Books.

Johnson, P. O. and Neyman, J. (1936). Tests of certain linear hypothesis and their application to some educational problems. *Statistical Research Memoirs*, 1, 57-93.

Johnson, P. O. and Fay, L. C. (1950). The Johnson-Neyman technique, its theory and application. *Pychometirika*, *15*, 349-367.

Huitema, B. E. (1980). *The analysis of covariance and alternatives*, New York, John Wiley & Sons.

Hunka, S. (1995). Identifying regions of significance in ANCOVA problems having non-homogeneous regressions. *British Journal of Mathematical and Statistical Psychology*, *48,* 161-188.

Hunka, S. & Legihton, J. (1997). Defining Johnson-Neyman regions of significance in the three-covariate ANCOVA using Mathematica. *Journal of Educational and Behavioral Statistics*, *Vol* 22, *No.* 4, 361-387.

Huttenlocher, J. E., Haight, W., Bryk, A. S., & Selzer, M. (1991). Early Vocabulary Growth: Relation to Language Input and Gender. *Developmental Psychology, 27*(2), 236-249.

Potthoff, R. F. (1964). On the Johnson-Neyman technique and some extensions thereof. *Psychometrika, 29*, 241-256.

Potthoff, R. F. (1983). Johnson-Neyman technique, In S. Kotz and N. L. Johnson (Eds.), *Encyclopedia of Statistical Sciences*, pp. 299-303. New York: Wiley.

Raudenbush, S. W., & Chan, W. S. (1993). Application of a hierarchical linear model to the study of adolescent deviant in an overlapping cohort design. *Journal of Consulting and Clinical Psychology. 61*(*6*), 941-951.

Raudenbush, S. W., Bryk, A., Cheong, Y. F., and Congdon, R. (2000). *HLM5 - Hierarchical Linear and Nonlinear Modeling*. Scientific Software International, Inc., Chicago, IL.

Wald, A. (1941). Asymptotically most powerful tests of statistical hypotheses. *Annals of Mathematical Statistics*, *12*, 1-19.

Wolfram, S. (1999). *The Mathematica Book. 4th Ed.* Wolfram Media/Cambridge University Press.

Figure 1.

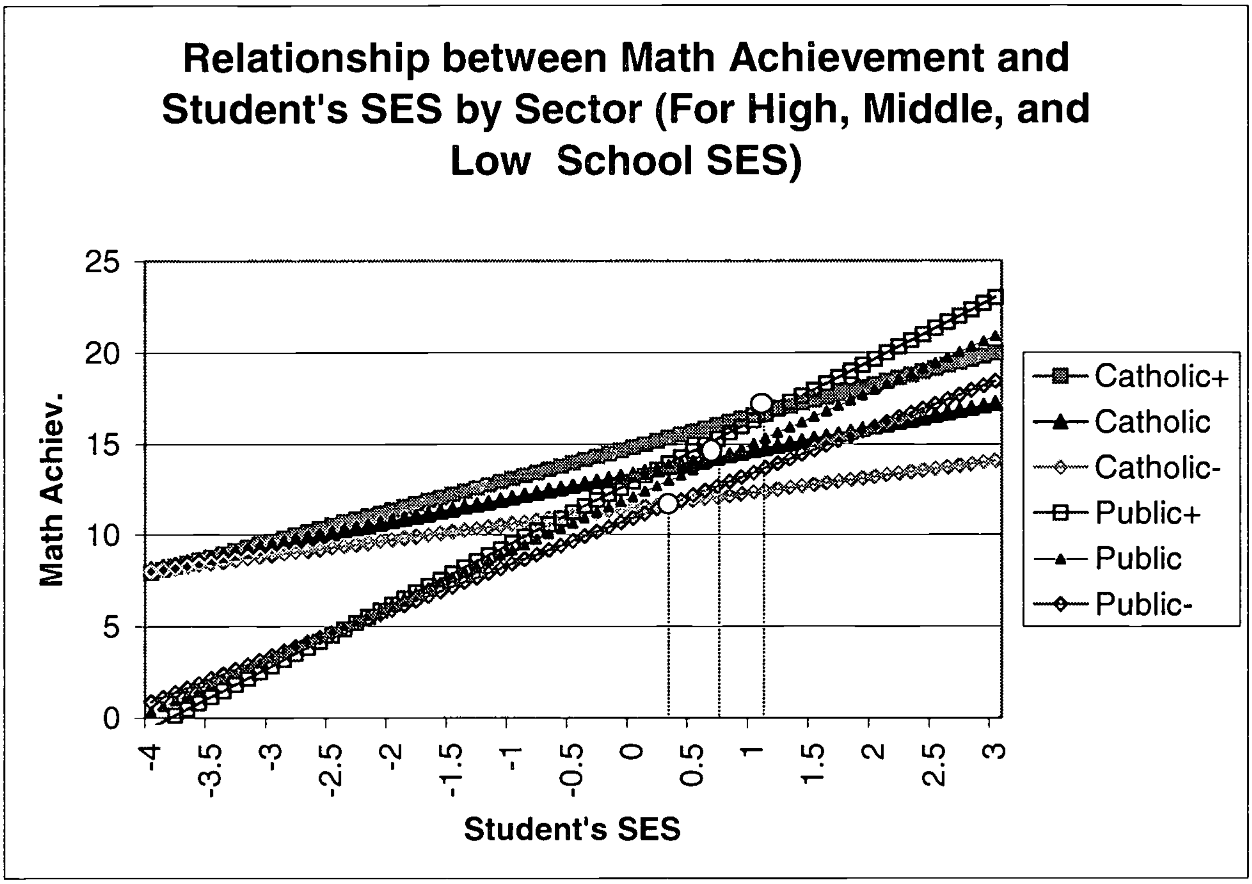
Relationship between Math Achievement and Student's SES by Sector (For High, Middle, and Low School SES)
25
20
15
10
5
0
Math Achiev.
-4
-3.5
-3
-2.5
-2
-1.5
-1
-0.5
0
0.5
1
1.5
2
2.5
3
Student's SES
Catholic+
Catholic
Catholic-
Public+
Public
Public-

Figure 2. A Plot of H – 3.84 as the Function of Relative Student SES

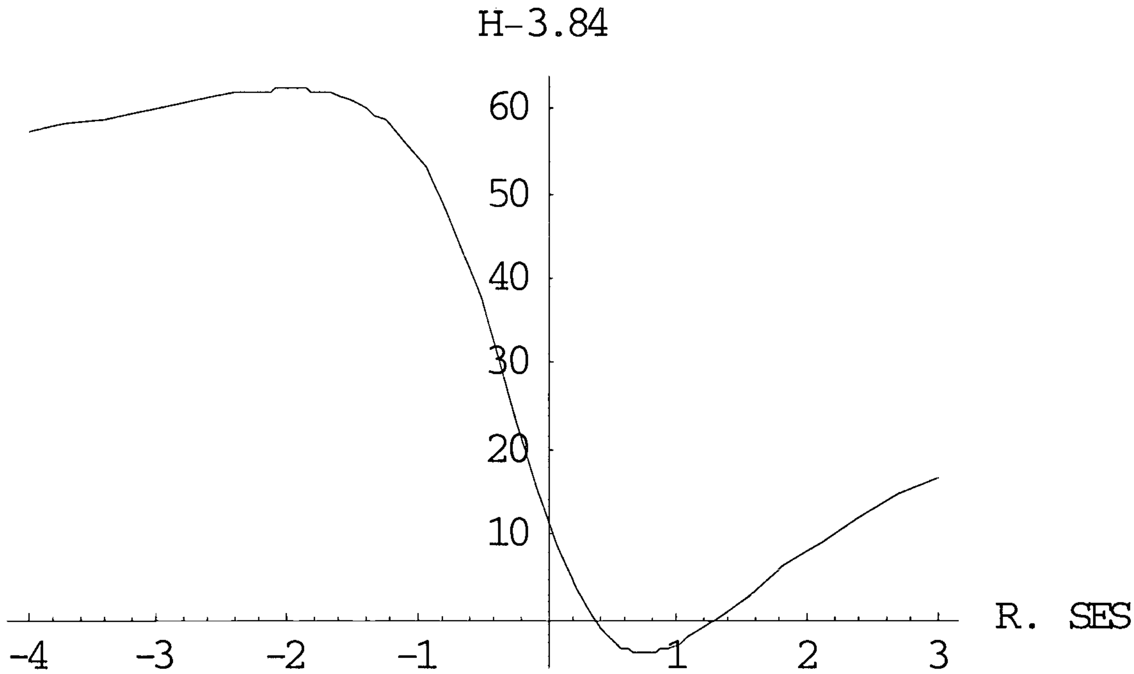

Figure 3. A 95 % Confidence Interval on the Difference in Math Achievement as the function of Relative Student's SES

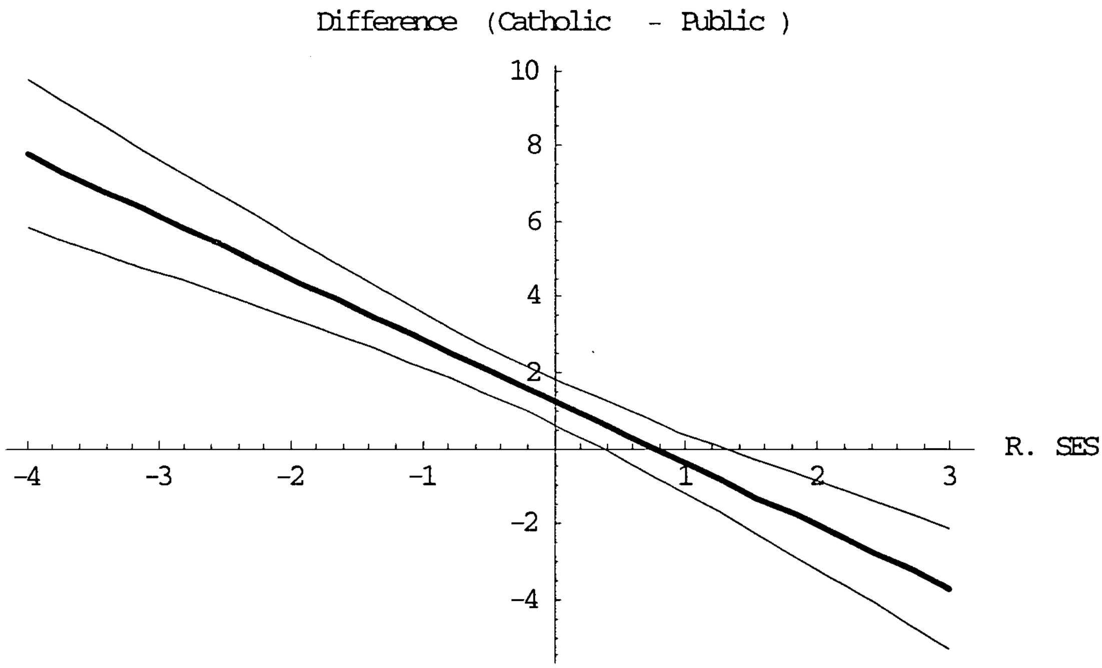

Figure 4. Region of Significant Difference on the Plane of Student's SES and School SES

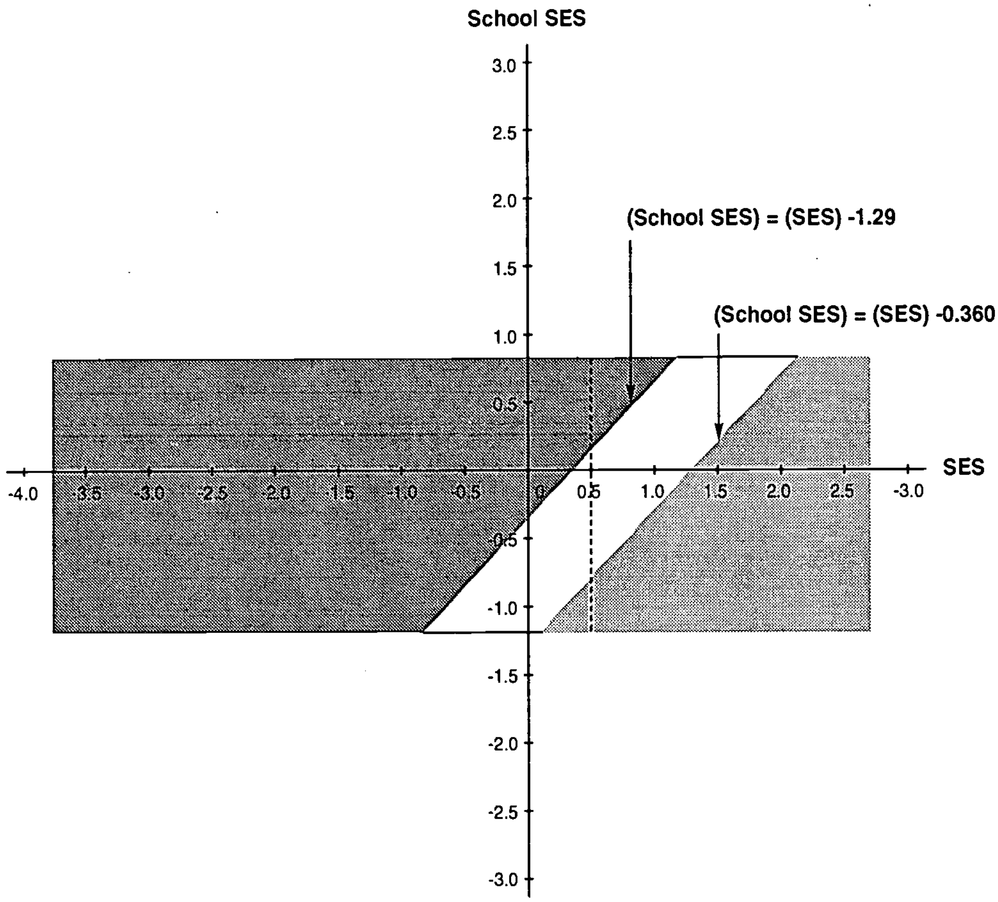

Figure 5.

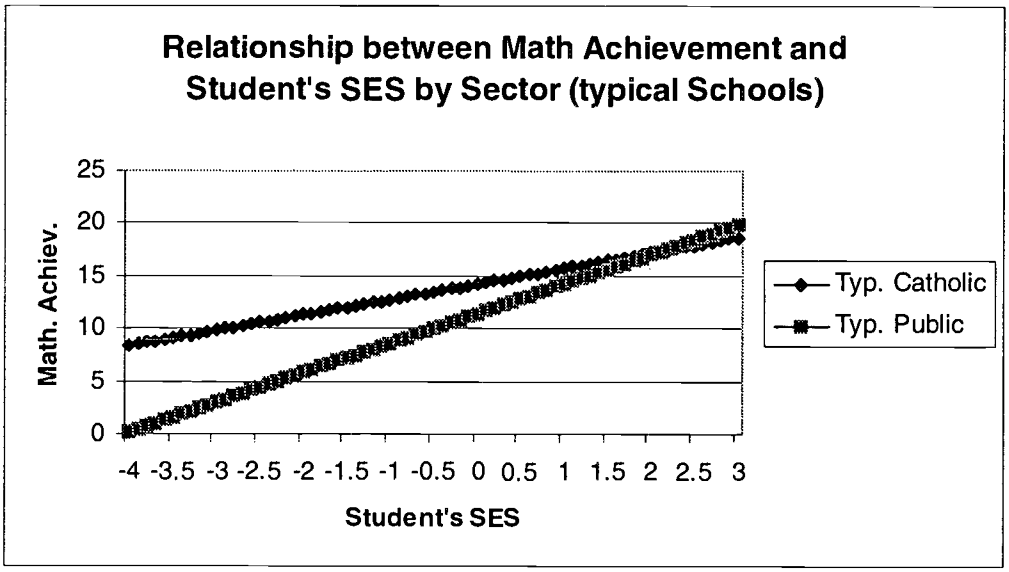
Relationship between Math Achievement and Student's SES by Sector (typical Schools)
Math. Achiev.
25
20
15
10
5
0
-4 -3.5 -3 -2.5 -2 -1.5 -1 -0.5 0 0.5 1 1.5 2 2.5 3
Student's SES
Typ. Catholic
Typ. Public

Figure 6.

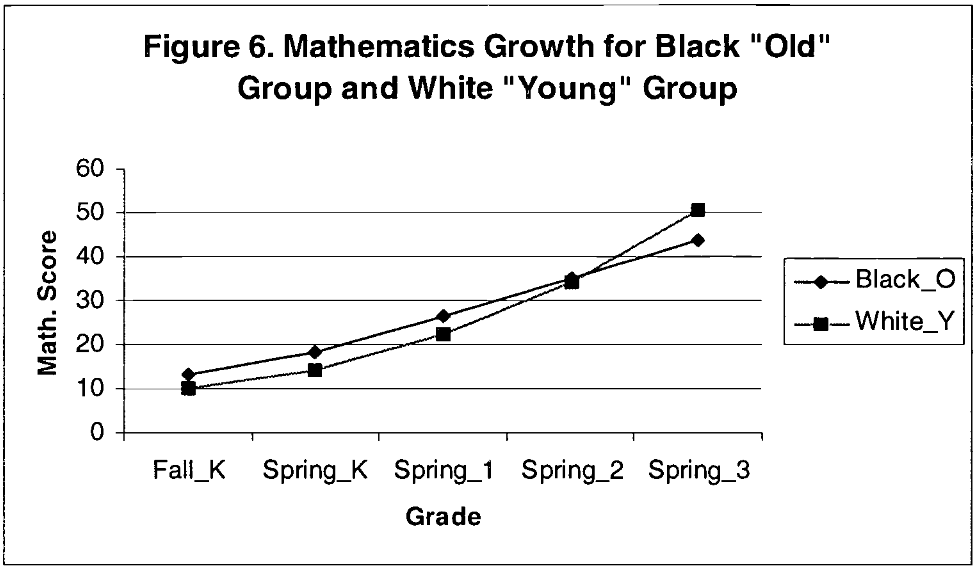
Figure 6. Mathematics Growth for Black "Old" Group and White "Young" Group
60
50
40
30
20
10
0
Math. Score
Fall_K
Spring_K
Spring_1
Spring_2
Spring_3
Grade
Black_O
White_Y

Figure 7. A Plot of H – 3.84 as the Function of Coded Grade

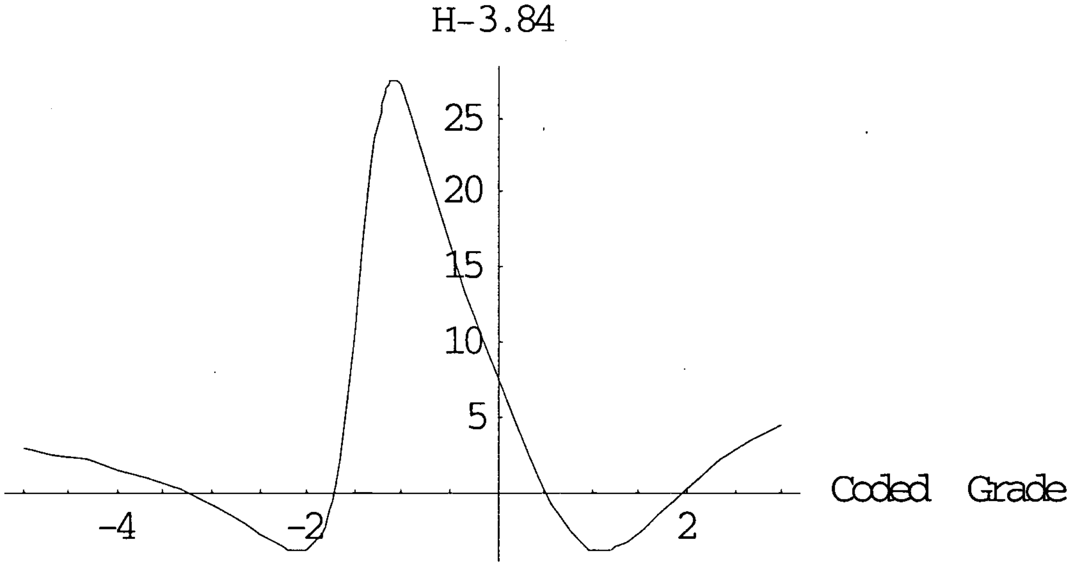

Figure 8. A Plot of the effect size $K^T\hat{\gamma}$ as the Function of Coded Grade

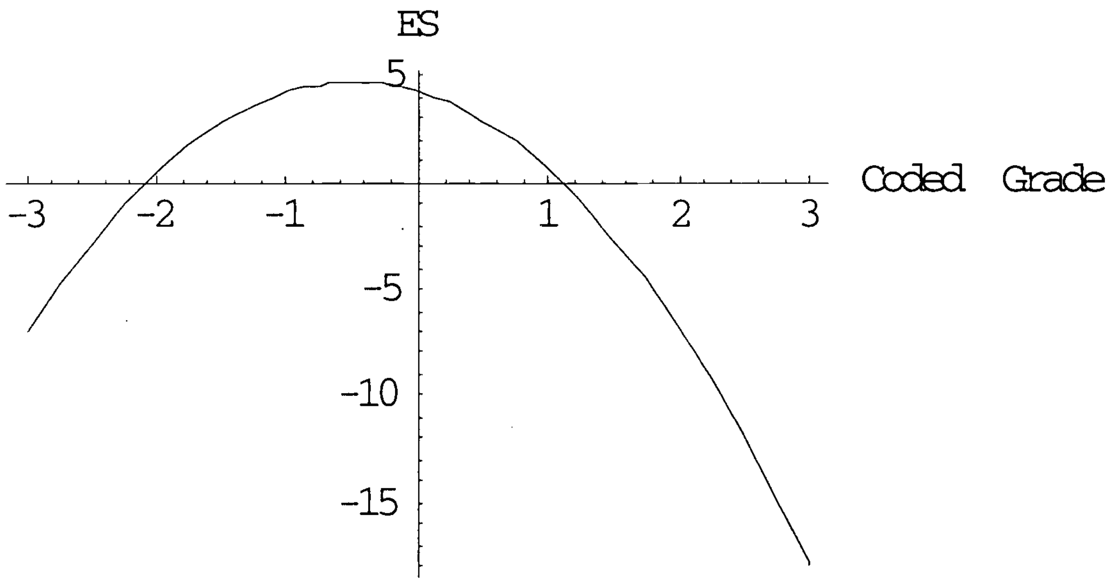

Figure 9. A Plot of the squae of the effect size $K^T\hat{\gamma}$ as the Function of Coded Grade

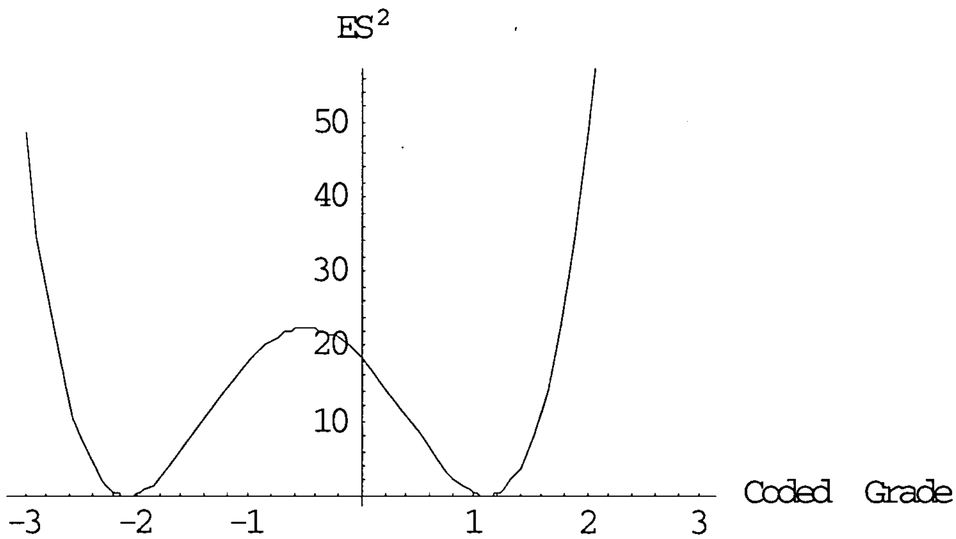

Figure 10. A Plot of the estimated Variance $\hat{V}_K$ as the Function of Coded Grade

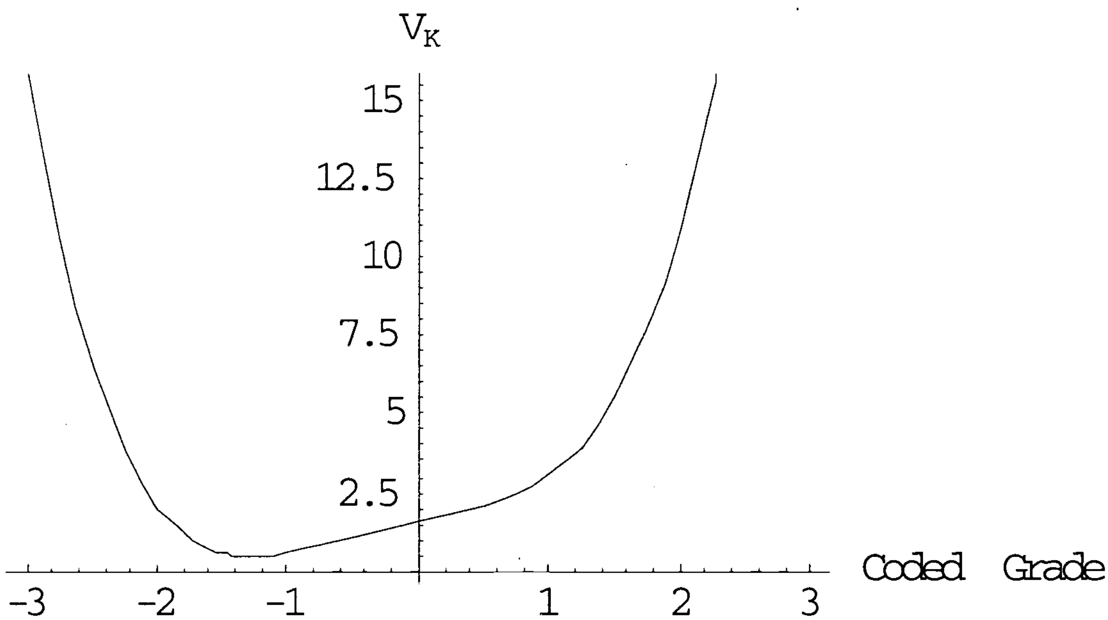

Table 1. Descriptive Statistics for High School & Beyond Data

Overall:

Level-1 Descriptive Statistics					
Variable Name	N	Mean	SD	Minimum	Maximum
MATHACH	7185	12.748	6.878	-2.832	24.993
SES	7185	0.000	0.779	-3.758	2.692
RSES	7185	0.000	0.661	-3.657	2.850

Level-2 Descriptive Statistics					
Variable Name	N	Mean	SD	Minimum	Maximum
MEANSES	160	0.000	0.414	-1.190	0.830
SECTOR	160	0.438	0.498	0	1

Note: 1. SECTOR = 1 if Catholic and = 0 if Public.
2. RSES is the relative student's SES, relative to the school mean SES.

Classified by SECTOR:

Catholic:

Level-1 Descriptive Statistics					
Variable Name	N	Mean	SD	Minimum	Maximum
MATHACH	3543	14.170	6.359	-2.832	24.993
SES	3543	0.150	0.741	-2.838	1.762
RSES	3543	0.000	0.623	-3.373	2.518

Level-2 Descriptive Statistics					
Variable Name	N	Mean	SD	Minimum	Maximum
MEANSES	70	0.166	0.396	-0.760	0.830

Public:

Level-1 Descriptive Statistics					
Variable Name	N	Mean	SD	Minimum	Maximum
MATHACH	3642	11.364	7.080	-2.832	24.993
SES	3642	-0.146	0.788	-3.758	2.692
RSES	3642	0.000	0.696	-3.657	2.857

Level-2 Descriptive Statistics					
Variable Name	N	Mean	SD	Minimum	Maximum
MEANSES	90	-0.130	0.383	-1.190	0.690

Table 2. Effect of Student and School SES on Mathematics Achievement

The outcome variable is MATHACH

Final estimation of fixed effects:

Fixed Effect	Coefficient	Standard Error	T-ratio	Approx. d.f.	P-value
For INTRCPT1, β_{0j}					
INTRCPT2, γ_{00}	12.095	0.199	60.865	157	0.000
MEANSES, γ_{01}	5.333	0.369	14.446	157	0.000
SECTOR, γ_{02}	1.226	0.306	4.004	157	0.000
For SES slope, β_{1j}					
INTRCPT2, γ_{10}	2.938	0.157	18.698	157	0.000
MEANSES, γ_{11}	1.034	0.303	3.419	157	0.001
SECTOR, γ_{12}	-1.641	0.243	-6.756	157	0.000

Final estimation of variance components:

Random Effect	Standard Deviation	Variance Component	df	Chi-square	P-value
INTRCPT1, u_{0j}	1.543	2.380	157	605.295	0.000
SES slope, u_{1j}	0.386	0.149	157	162.309	0.369
level-1, ε_{ij}	6.058	36.703			

Table 3. Classification of the students by sector and level of their SES

	RSES < 0.36	0.36 ≤ RSES < 1.29	1.29 ≤ RSES	Total
Catholic	2469 (69.6%)	1039 (29.3%)	35 (1.0%)	3543 (100%)
Public	2524 (69.3%)	1025 (28.1%)	93 (2.6%)	3642 (100%)
Total	4993	2064	128	7185

Table 4. Descriptive Statistics for Early Child Cognitive Development Data

a) Mathematics Score (Outcome Variable)

Overall:

	N	Mean	Std. Deviation	Minimum	Maximum
MATH1	493	11.67	5.53	1	30
MATH2	484	16.44	6.69	3	42
MATH3	372	25.57	10.11	4	58
MATH4	284	38.13	12.83	9	71
MATH5	92	50.86	13.01	19	74

Classified by 6 Groups:

Descriptives

		N	Mean	Std. Deviation	Minimum	Maximum
MATH1	White_Old	104	16.35	5.84	6	30
	White_Middle	100	12.95	4.85	4	25
	White_Young	50	11.74	4.69	1	25
	Balck_Old	69	10.39	4.91	1	28
	Black_Middle	87	8.95	3.94	3	23
	Black_Young	83	8.14	3.35	2	18
	Total	493	11.67	5.53	1	30
MATH2	White_Old	101	22.18	6.46	7	42
	White_Middle	100	18.03	5.65	6	30
	White_Young	49	16.39	5.39	5	29
	Balck_Old	68	14.57	5.00	3	26
	Black_Middle	87	13.75	6.07	4	28
	Black_Young	79	11.67	4.82	3	34
	Total	484	16.44	6.69	3	42
MATH3	White_Old	95	32.37	10.29	7	58
	White_Middle	77	27.65	9.00	9	55
	White_Young	40	26.03	10.62	9	55
	Balck_Old	49	22.14	9.25	7	53
	Black_Middle	53	20.98	5.77	8	34
	Black_Young	58	18.45	6.23	4	39
	Total	372	25.57	10.11	4	58
MATH4	White_Old	81	45.44	10.62	20	71
	White_Middle	65	42.66	12.05	17	66
	White_Young	30	37.80	11.84	19	64
	Balck_Old	27	32.52	14.38	9	61
	Black_Middle	39	30.51	9.68	15	59
	Black_Young	42	27.93	7.16	13	45
	Total	284	38.13	12.83	9	71
MATH5	White_Old	42	57.71	8.66	39	74
	White_Middle	0	.	.	.	.
	White_Young	14	54.71	8.29	40	74
	Balck_Old	13	41.77	13.24	22	59
	Black_Middle	0	.	.	.	.
	Black_Young	23	41.13	13.30	19	63
	Total	92	50.86	13.01	19	74

b) Covariates:

Overall:

	N	Mean	Std. Deviation	Minimum	Maximum
IQ	493	97.50	15.27	52	141
LIT	493	9.96	3.78	.00	17.00
MOMED	493	13.64	2.39	6	23
DO	493	.35	.48	0	1
DM	493	.38	.49	0	1

Classified by 6 groups:

Descriptives

		N	Mean	Std. Deviation	Minimum	Maximum
IQ	White_Old	104	106.95	14.85	72	134
	White_Middle	100	103.90	12.89	75	134
	White_Young	50	107.24	12.61	82	141
	Balck_Old	69	88.90	13.16	52	126
	Black_Middle	87	87.18	11.04	53	121
	Black_Young	83	90.05	10.10	66	122
	Total	493	97.50	15.27	52	141
LIT	White_Old	104	12.90	2.58	5.00	17.00
	White_Middle	100	12.11	2.89	2.00	17.00
	White_Young	50	11.99	2.26	6.00	15.00
	Balck_Old	69	7.34	2.90	.00	17.00
	Black_Middle	87	6.85	2.75	1.18	15.00
	Black_Young	83	7.88	3.13	2.61	16.00
	Total	493	9.96	3.78	.00	17.00
MOMED	White_Old	104	14.93	2.32	10	21
	White_Middle	100	14.39	2.40	8	18
	White_Young	50	14.02	2.36	8	20
	Balck_Old	69	12.48	2.17	6	18
	Black_Middle	87	12.70	1.71	6	18
	Black_Young	83	12.86	2.20	8	23
	Total	493	13.64	2.39	6	23

c) HLM Data Set Information

Level-1 Descriptive Statistics					
Variable Name	N	Mean	SD	Minimum	Maximum
MATH	1725	22.45	14.43	1.00	74.00
GRADE	1725	-0.44	1.05	-1.50	2.00
GRADESQ	1725	1.30	1.03	0.00	4.00
DFALL	1725	0.29	0.45	0.00	1.00

Level-2 Descriptive Statistics					
Variable Name	N	Mean	SD	Minimum	Maximum
DBLACK	493	0.48	0.50	0.00	1.00
IQ	493	97.50	15.27	52.00	141.00
LIT	493	9.96	3.78	0.00	17.00
MOMED	493	13.64	2.39	6.00	23.00
DO	493	0.35	0.48	0.00	1.00
DM	493	0.38	0.49	0.00	1.00

Table 5. Child Cognitive Development

The outcome variable is MATH

Final estimation of fixed effects:

Fixed Effect	Coefficient	Standard Error	T-ratio	Approx. d.f.	P-value
For INTRCPT1, β_{0j}					
INTRCPT2, γ_{00}	22.179	0.795	27.903	486	0.000
DBLACK, γ_{01}	-1.410	0.825	-1.710	486	0.087
IQ, γ_{02}	0.346	0.026	13.288	486	0.000
LIT, γ_{03}	0.153	0.069	2.208	486	0.027
DO, γ_{04}	5.694	0.826	6.894	486	0.000
DM, γ_{05}	3.404	0.834	4.079	486	0.000
MOMED, γ_{06}	0.215	0.087	2.458	486	0.014
For GRADE slope, β_{1j}					
INTRCPT2, γ_{10}	10.115	0.481	21.034	488	0.000
DBLACK, γ_{11}	-1.860	0.486	-3.825	488	0.000
IQ, γ_{12}	0.077	0.015	5.059	488	0.000
DO, γ_{13}	0.029	0.481	0.060	488	0.953
DM, γ_{14}	0.980	0.527	1.861	488	0.062
For GRADESQ slope, β_{2j}					
INTRCPT2, γ_{20}	2.004	0.370	5.413	488	0.000
DBLACK, γ_{21}	-0.811	0.374	-2.169	488	0.030
IQ, γ_{22}	-0.029	0.012	-2.492	488	0.013
DO, γ_{23}	-1.051	0.354	-2.969	488	0.003
DM, γ_{24}	-0.157	0.425	-0.370	488	0.711
For DFALL slope, β_{3j}					
INTRCPT2, γ_{30}	-1.380	0.473	-2.915	1707	0.004

Final estimation of variance components:

Random Effect	Standard Deviation	Variance Component	df	Chi-square	P-value
INTRCPT1, u_{0j}	5.416	29.330	363	1135.791	0.000
GRADE slope, u_{1j}	2.503	6.265	365	729.976	0.000
GRADESQ slope, u_{2j}	1.275	1.626	365	442.752	0.003
level-1, ε_{ij}	4.437	19.690			

Table 6. Predicted Differences between Black "Old" and White "Young" Groups at each Grade and their Chi-Square Tests

Grade	Code	Black_Old	White_Young	Difference	Chi-Square	P-Value
Fall, K	-1.5	12.98	10.14	2.84	13.85	0.0004
Spring, K	-1.0	18.32	14.07	4.25	30.79	0.000006
Spring, 1st	0.0	26.46	22.18	4.28	11.60	0.001
Spring, 2nd	1.0	34.89	34.30	0.59	0.12	>0.5
Spring, 3rd	2.0	43.60	50.43	-6.83	4.29	0.036

U.S. Department of Education
Office of Educational Research and Improvement (OERI)
National Library of Education (NLE)
Educational Resources Information Center (ERIC)

ERIC®

REPRODUCTION RELEASE

TM033861

(Specific Document)

I. DOCUMENT IDENTIFICATION:

Title: Johnson-Neyman type Technique in Hierarchical Linear Model	
Author(s): Yasuo Miyazaki	
Corporate Source: University of Michigan	Publication Date: 4/1/02

II. REPRODUCTION RELEASE:

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

If permission is granted to reproduce and disseminate the identified document, please CHECK ONE of the following three options and sign at the bottom of the page.

The sample sticker shown below will be affixed to all Level 1 documents	The sample sticker shown below will be affixed to all Level 2A documents	The sample sticker shown below will be affixed to all Level 2B documents
PERMISSION TO REPRODUCE AND DISSEMINATE THIS MATERIAL HAS BEEN GRANTED BY Sample TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC) 1	PERMISSION TO REPRODUCE AND DISSEMINATE THIS MATERIAL IN MICROFICHE, AND IN ELECTRONIC MEDIA FOR ERIC COLLECTION SUBSCRIBERS ONLY, HAS BEEN GRANTED BY Sample TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC) 2A	PERMISSION TO REPRODUCE AND DISSEMINATE THIS MATERIAL IN MICROFICHE ONLY HAS BEEN GRANTED BY Sample TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC) 2B
Level 1	Level 2A	Level 2B
☑	☐	☐
Check here for Level 1 release, permitting reproduction and dissemination in microfiche or other ERIC archival media (e.g., electronic) *and* paper copy.	Check here for Level 2A release, permitting reproduction and dissemination in microfiche and in electronic media for ERIC archival collection subscribers only	Check here for Level 2B release, permitting reproduction and dissemination in microfiche only

Documents will be processed as indicated provided reproduction quality permits.
If permission to reproduce is granted, but no box is checked, documents will be processed at Level 1.

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

Sign here,→ please

Signature: Yasuo Miyazaki	Printed Name/Position/Title: Yasuo Miyazaki, Assist. Research Scientist	
Organization/Address: Center for Human Growth & Development 300 N. Ingalls Building, 10th level Ann Arbor, MI 48109	Telephone: 734-764-2443	FAX: 734-936-9288
	E-Mail Address: yasuom@umich.edu	Date: 4/3/02

(over)

III. DOCUMENT AVAILABILITY INFORMATION (FROM NON-ERIC SOURCE):

If permission to reproduce is not granted to ERIC, *or*, if you wish ERIC to cite the availability of the document from another source, please provide the following information regarding the availability of the document. (ERIC will not announce a document unless it is publicly available, and a dependable source can be specified. Contributors should also be aware that ERIC selection criteria are significantly more stringent for documents that cannot be made available through EDRS.)

Publisher/Distributor:
Address:
Price:

IV. REFERRAL OF ERIC TO COPYRIGHT/REPRODUCTION RIGHTS HOLDER:

If the right to grant this reproduction release is held by someone other than the addressee, please provide the appropriate name and address:

Name:
Address:

V. WHERE TO SEND THIS FORM:

Send this form to the following ERIC Clearinghouse:

University of Maryland
ERIC Clearinghouse on Assessment and Evaluation
1129 Shriver Laboratory
College Park, MD 20742
Attn: Acquisitions

However, if solicited by the ERIC Facility, or if making an unsolicited contribution to ERIC, return this form (and the document being contributed) to:

ERIC Processing and Reference Facility
4483-A Forbes Boulevard
Lanham, Maryland 20706

Telephone: 301-552-4200
Toll Free: 800-799-3742
FAX: 301-552-4700
e-mail: ericfac@inet.ed.gov
WWW: http://ericfac.piccard.csc.com

EFF-088 (Rev. 2/2000)