Do Green Schoolyards Grow Learning?

Steven Lin
Manny Yu
Ming Sun

STAT 427
Spring 2013

ABSTRACT

The purpose of this project was to investigate the effect of green schoolyards on academic achievement using an objective greenness measure. More specifically, tree canopy area was used as the basis for the different greenness measures considered in the study. The scope of the project was the elementary and middle grades in public schools in Chicago, IL. The client for the project was Dr. Kuo from the University of Illinois at Urbana-Champaign.

The data preparation was a significant portion of the project and is discussed in detail. Descriptive statistics at the school and student level are also provided. Hierarchical linear model was used to analyze the greenness effect on academic achievement. In addition, as requested by the client, step-by-step instructions to replicate the analysis in this project are provided in the appendix.

CONTENTS

LIST OF TABLES

LIST OF FIGURES

1. INTRODUCTION

1.1 Motivation

The motivation of the project was to determine if schoolyard greenness has an impact on academic achievement. The hypothesis is that green schoolyard might enhance attention, which is a key factor in learning.

1.2 Previous study

The scope of the project was the elementary and middle grades in public schools in Chicago, IL. The previous study in this area used subjective 5-point ratings of aerial photos as a greenness measure of the schoolyard. An example of the ratings is shown below in Figure 1. The subjective ratings are limited because the ratings are based on human subjective evaluations, and thus the ratings might be difficult to replicate and are subject to bias. Another limitation is the discrete and limited nature of the 5-point rating scale, which restricts the differentiation of schoolyards in terms of greenness.

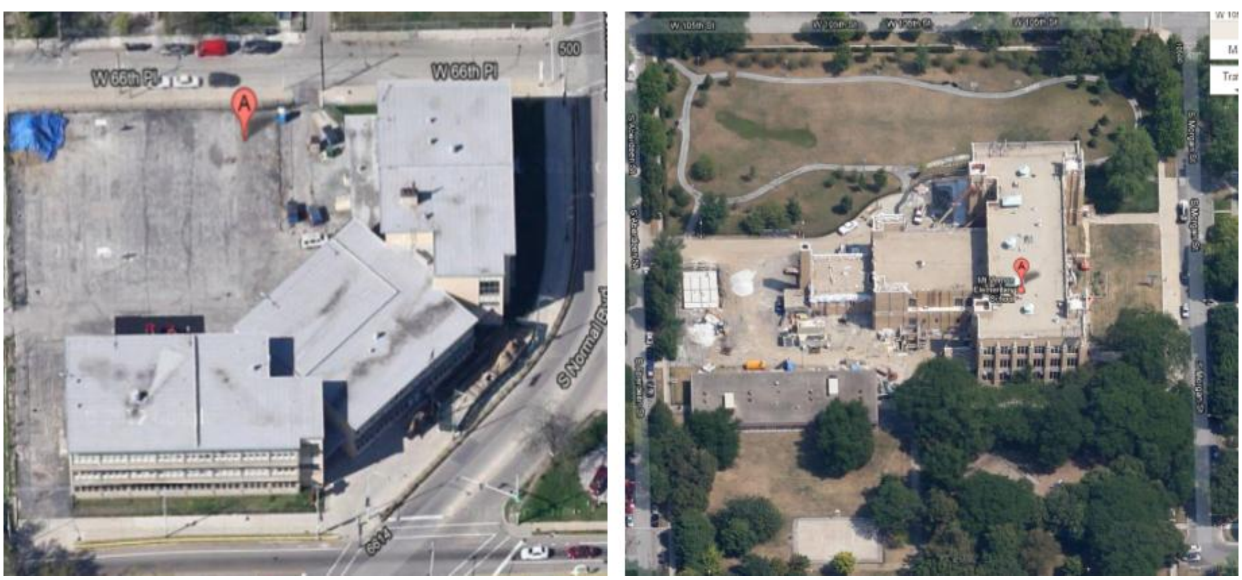


Figure 1. Subjective 5-point ratings. Left: low greenness (rating =1). Right: high greenness (rating =5)

1.3 Current Study

The current study uses tree canopy coverage area as a greenness measure. Tree canopy is defined as the foliage of trees that cover the ground when viewed from above. This is graphically illustrated in Figure 2. Compared to the subjective 5-point rating, the tree canopy area addresses the limitations since it is objective and continuous (in units of ft^2). The measurement process is shown in Figure 3. The sun emits incoming electromagnetic radiation (EMR), which is reflected from the surface of objects in earth, and then captured by the satellite. For trees, the satellite is able to measure the canopy area by capturing a specific wavelength: the photosynthesis of plants converts incoming EMR into energy and reflects it back at near infrared (IR) wavelength.

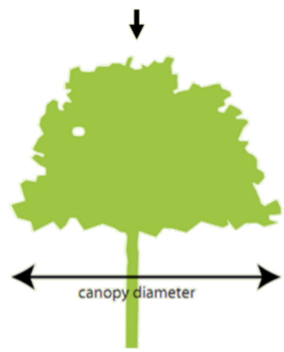


Figure 2. Tree canopy

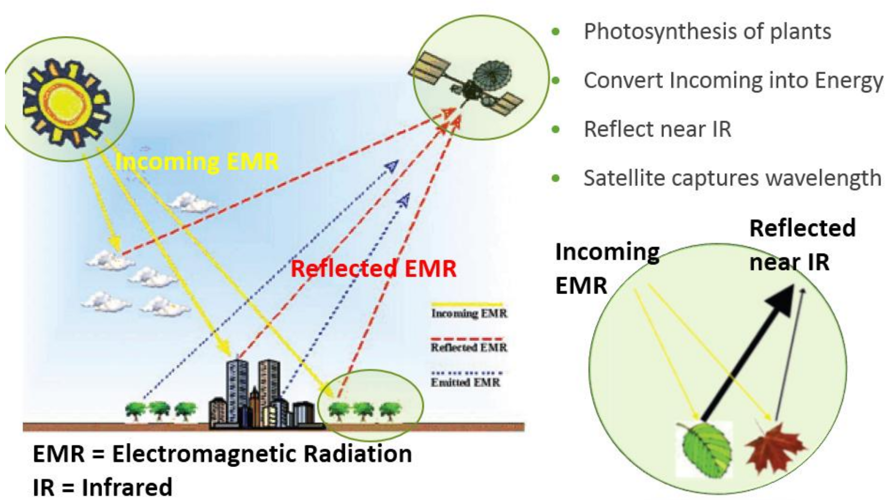


Figure 3. Tree canopy area measurement process

1.4 GIS Dataset

The tree canopy measurements were recorded in a GIS (geographic information system) format and provided by a consultant for the city of Chicago. For this project, the software ArcGIS, a mapping and spatial analysis tool, was used extensively to understand the data and prepare it for analysis.

The graphical representation of the GIS dataset is shown in Figure 4. GIS data basically breaks down the 3D real world into various 2D layers such as land usage, elevation parcel, streets or customers. Each layer is composed of spatial data (e.g. X and Y coordinates) and attribute data (e.g. school name, property area, tree canopy area, landscape area and building area). Thus, each row in the GIS dataset refers to a specific school with its attribute data information. The spatial data is graphically illustrated in Figure 5, and as expected, the number of greener schools increases as the distance from the center of the city increases. The important attribute data (e.g. school property area, tree canopy area and landscape area) is also shown graphically in Figure 6.

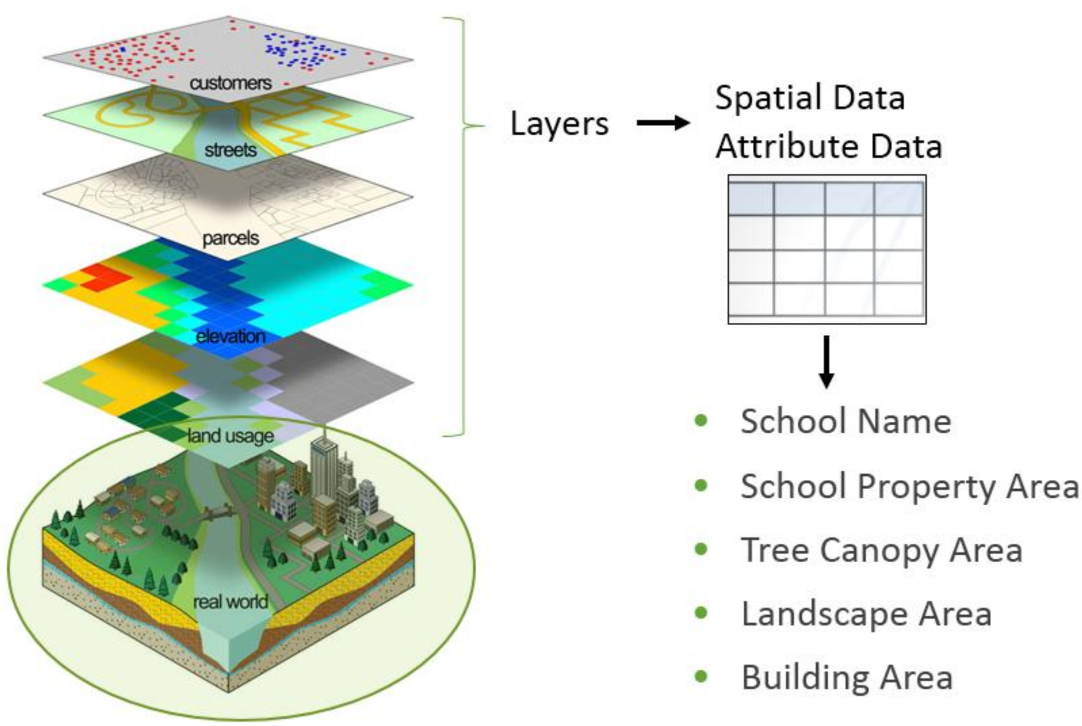


Figure 4. GIS dataset

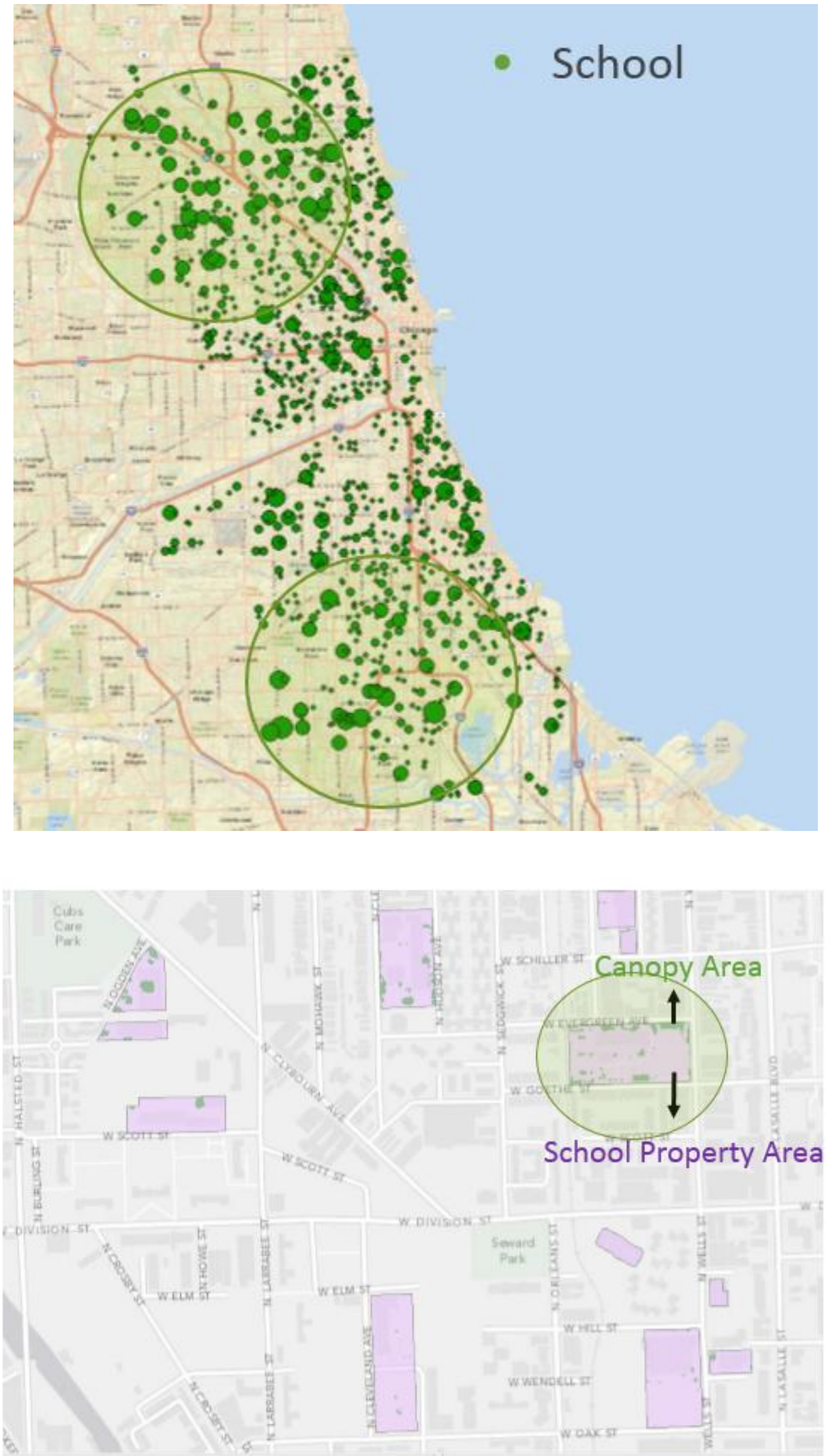


Figure 5. GIS spatial data

Figure 6. GIS attribute data

2. DATA PREPARATION

2.1 Matching

Data preparation was a significant part of the project. One of the reasons was that the GIS dataset did not contain any information on the variables and the names of the variables were not descriptive. Thus, ArcGIS was used extensively to deduce the meaning of the variables provided in the GIS dataset.

The first step was to match the GIS dataset to the school dataset to the school-level dataset. The matching process is depicted below in Figure 7. The GIS dataset contains 958 schools, while the school dataset contains 444 schools, which include only elementary and middle grades in public schools. Because the school names of the GIS dataset and school dataset were not exactly the same, a school ID was created to link both datasets. The resulting school-level dataset contained 438 data points (i.e. schools).

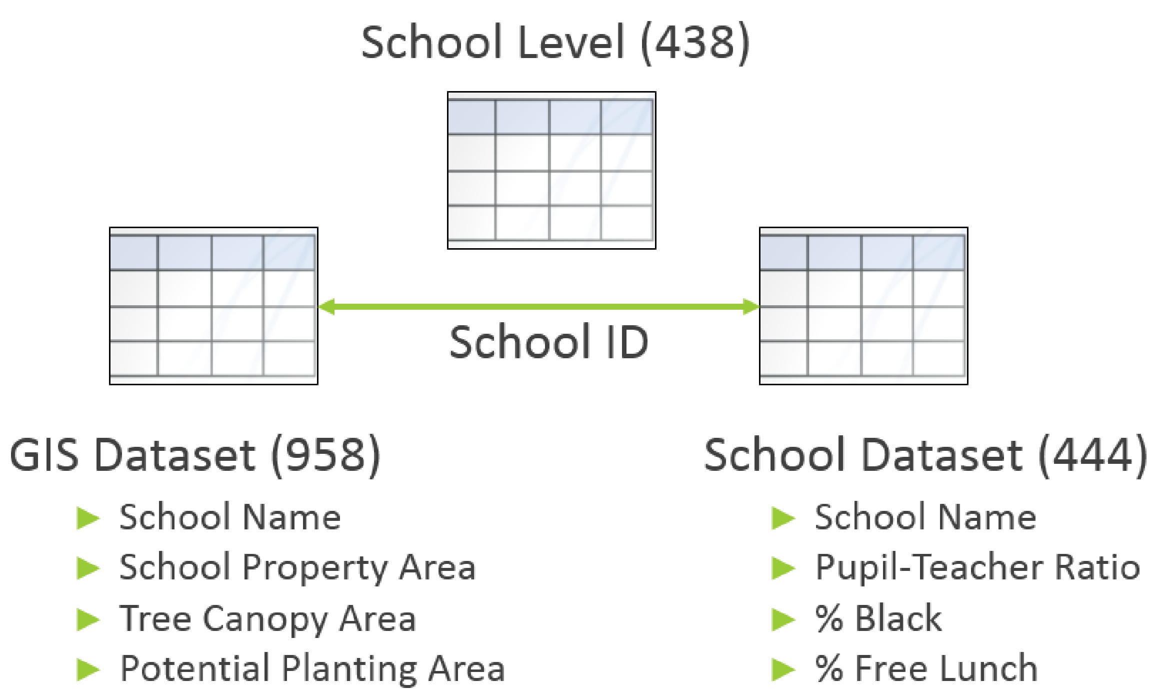


Figure 7. Matching datasets

2.2 Building Area

Originally, the building information for the schools was not provided in the dataset. Because the greenness measure of the schoolyard might depend on the building size, all the building footprints in Chicago (820,154) were obtained from the city of Chicago official website[1]. The objective was to have a building area for each school. However, there was no matching variable in the building dataset and school-level dataset. Therefore, the intersect function in ArcGIS was used to link the school to its corresponding building area. Figure 8 shows how the intersect function in ArcGIS works.

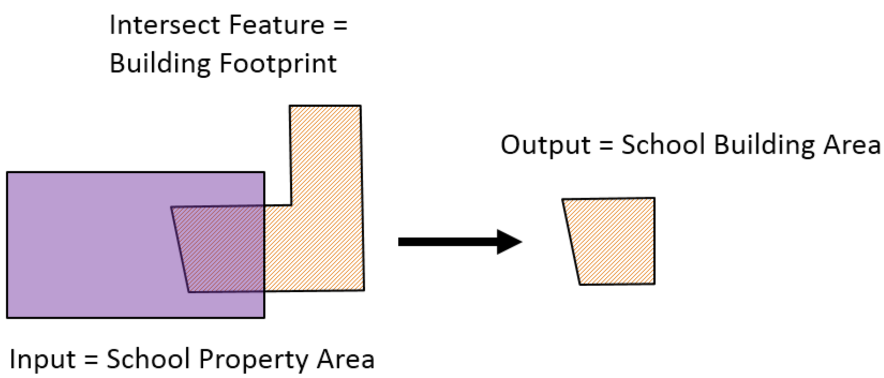


Figure 8. Intersect function in ArcGIS

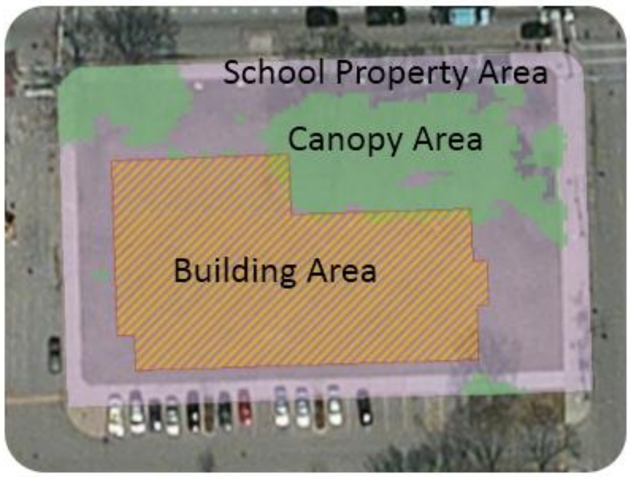
2.3 Greenness Measures

The canopy area is an objective measure of the greenness but by itself does not take into account the size of the school. Thus using just the canopy area might not be an accurate reflection of the greenness of the schoolyard. For example, a school with the same canopy but twice the school property area should be "less green" because the canopy is more scattered.

Thus, in addition to the canopy area, three more GIS-based greenness measures were considered that might be better at capturing the greenness of the schoolyard. A graphical summary of the greenness measures is provided in Figure 9. The ratio of canopy area to school property area was used because is a measure of the greenness per square foot of the school

[1] http://www.cityofchicago.org/

property. Similarly, the ratio of canopy area to school building area was used because is a measure of the greenness per square foot of the school building. The ratio of canopy area to potential planting area (PPA) was suggested to be a greenness measure and was originally provided with the GIS dataset. Although this ratio did not seem to be reliable after manually checking a few data points, the ratio was considered for further investigation.



- CanArea = Canopy Area
- CanPro_r = Canopy Area / School Property Area
- CanBld_r = Canopy Area / School Building Area
- CanPPA_r= Canopy Area / PPA

Figure 9. Greenness measures

2.4 Problematic Points

It is expected that the subjective rating and GIS-based greenness measures to be correlated. In order to compare the subjective ratings to the new greenness measures, each continuous measure was discretized into a 5-point rating scale based on the distribution of the subjective ratings (Figure 10). Therefore, if large differences exist between the subjective ratings and objective ratings, then this might be an indication of possible problematic data points that need further examination.

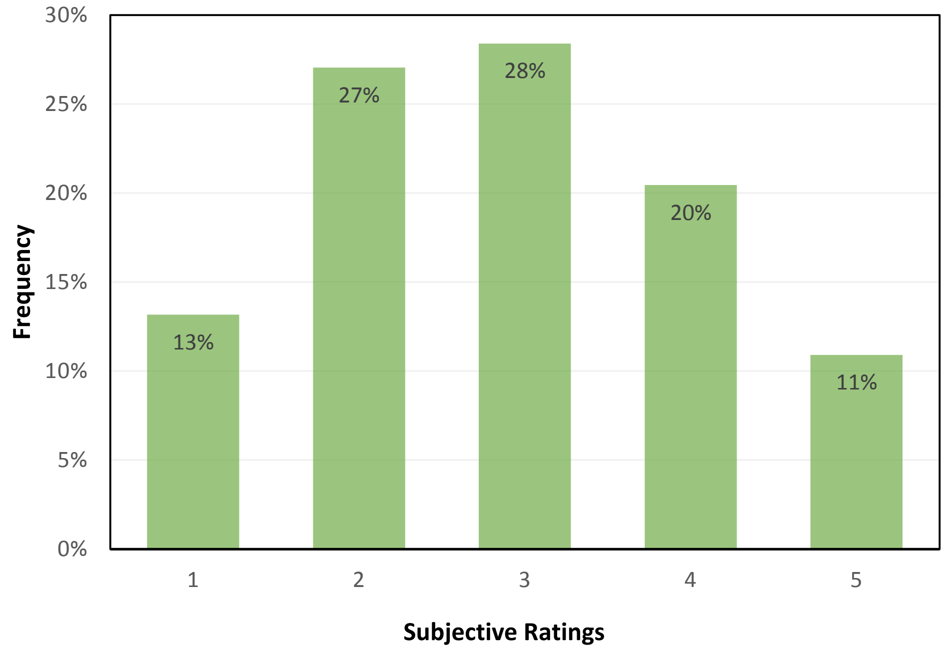


Figure 10. Distribution of subjective ratings

Cross-classification tables of the subjective ratings and each objective rating were constructed to identify problematic points. For example, Table 1 shows the cross-tabulation of the subjective rating and the objective rating based on canopy area.

Table 1. Ratings cross-tabulation (subjective vs. canopy area)

		Objective Rating				
		1	2	3	4	5
Subjective Rating	1	22	30	3	2	0
	2	15	51	40	13	0
	3	6	24	54	35	6
	4	9	11	22	32	15
	5	5	3	6	7	27

The table indicates that most point seem to fall along the diagonal, suggesting a general agreement. However, the points highlighted in red were identified as problematic points due the large discrepancy between the ratings. It can also be noticed that most of these points have a subjective ratings greater than the objective ratings. In other words, most problematic points occur when the school was rated as a high green school, but is a low green school according to the GIS data. More specifically, a problematic point was defined as point with an absolute difference between the objective rating and subjective rating of greater than 3 in at least on GIS-based greenness measure (e.g. canopy area, canopy area/property area, canopy area/building area). Under this criterion, a total of 27 data points (i.e. schools) were identified.

One example is Pershing Magnet (Figure 1), which had a subjective rating of 5 and objective ratings of 5. As it can be seen, this is clearly due to the fact that the property area is basically the building area. According to GIS, the schoolyard and visible greenness are not part of the school property.

Figure 11. Pershing Magnet

Based on this observation, a systematic rule to remove problematic points was created. The data points that had an absolute difference between the property area and building area of less than 100 ft^2 (i.e. schools where the property area is the building area) were removed. The reason was that the schoolyard was not part of the school property according to GIS. Under this rule, 12 data points were removed, resulting in a final school-level dataset of 420 data points. In addition, the analysis in this section identified points where the property area was less than the building area. These points were manually adjusted with the correct building area.

2.5 Limitations of GIS data

The analysis conducted during the data preparation stage was crucial in gaining a better understanding of the data, which made it possible to create GIS-based greenness measures and identify problematic points in the GIS dataset.

The analysis also pointed out a few limitations in the GIS dataset that were not apparent at first. One limitation is that the GIS data is restricted to information inside the school property as defined in GIS. Therefore, visible greenness and the schoolyard might not be captured by GIS if they fall outside the school property area. Additional limitations include the the fact that GIS ignores the distance between trees and buildings, and that no turf information is available. There is also a 5-year gap between the GIS data and the school data, but this is not expected to be a problem since most schools do not change significantly in a 5-year period of time. A

3. DESCRIPTIVE STATISTICS

3.1 School-Level Data

The school-level data contains 420 schools, and includes the greenness measures and school-specific variables. Table 2 provides a summary of the school-level variables.

Table 2. School-level variables

Variable Name	Description	Values
schno	School ID	Numeric
Pfree	% Free Lunch	Numeric
Puptch	Pupil-Teacher Ratio	Numeric
Pblack	% Percent Black	Numeric
CanArea	Canopy Area	Numeric
CanPro_r	Canopy Area / School Property Area	Numeric
CanBld_r	Canopy Area / School Building Area	Numeric
CanPPA_r	Canopy Area / PPA	Numeric

Not all the variables that measure the schools' greenness derived from GIS data might be appropriate. Although the objective measures and subjective ratings represent are different, it is expected that they will be consistent with each other. Thus, correlation analysis between the old ratings and the four greenness measurements derived from GIS data was conducted and is shown below in Table 3.

Table 3. Correlation between subjective ratings and GIS-based greenness measurements

Pearson Correlation Coefficients, N = 420 Prob > |r| under H0: Rho=0				
	Canopy Area	**Canopy/Property**	**Canopy/Building**	**Canopy/PPA**
Subjective Rating	**0.53322** **<.0001**	**0.55584** **<.0001**	**0.52860** **<.0001**	**0.03896** **0.4226**
Spearman Correlation Coefficients, N = 420 **Prob > |r| under H0: Rho=0**				
	Canopy Area	**Canopy/Property**	**Canopy/Building**	**Canopy/PPA**
Subjective Rating	**0.61644** **<.0001**	**0.55497** **<.0001**	**0.57431** **<.0001**	**0.10660** **0.0278**

Both the Pearson Correlation and Spearman Correlation suggest that ratio between canopy area and potential planting area is not reliable since it is poorly related with the subjective ratings. Thus, the other three GIS-based greenness measures that are highly positive correlated with the subjective ratings were included in the analysis.

Since the essence of the hierarchical linear model is linear model, the collinearity between the school-level variables may confound their effects when using them to predict the coefficients of the student-level variables. The collinearity problem was check by first looking at the scatterplot (Figure 12) and correlation heat map (Figure 13).

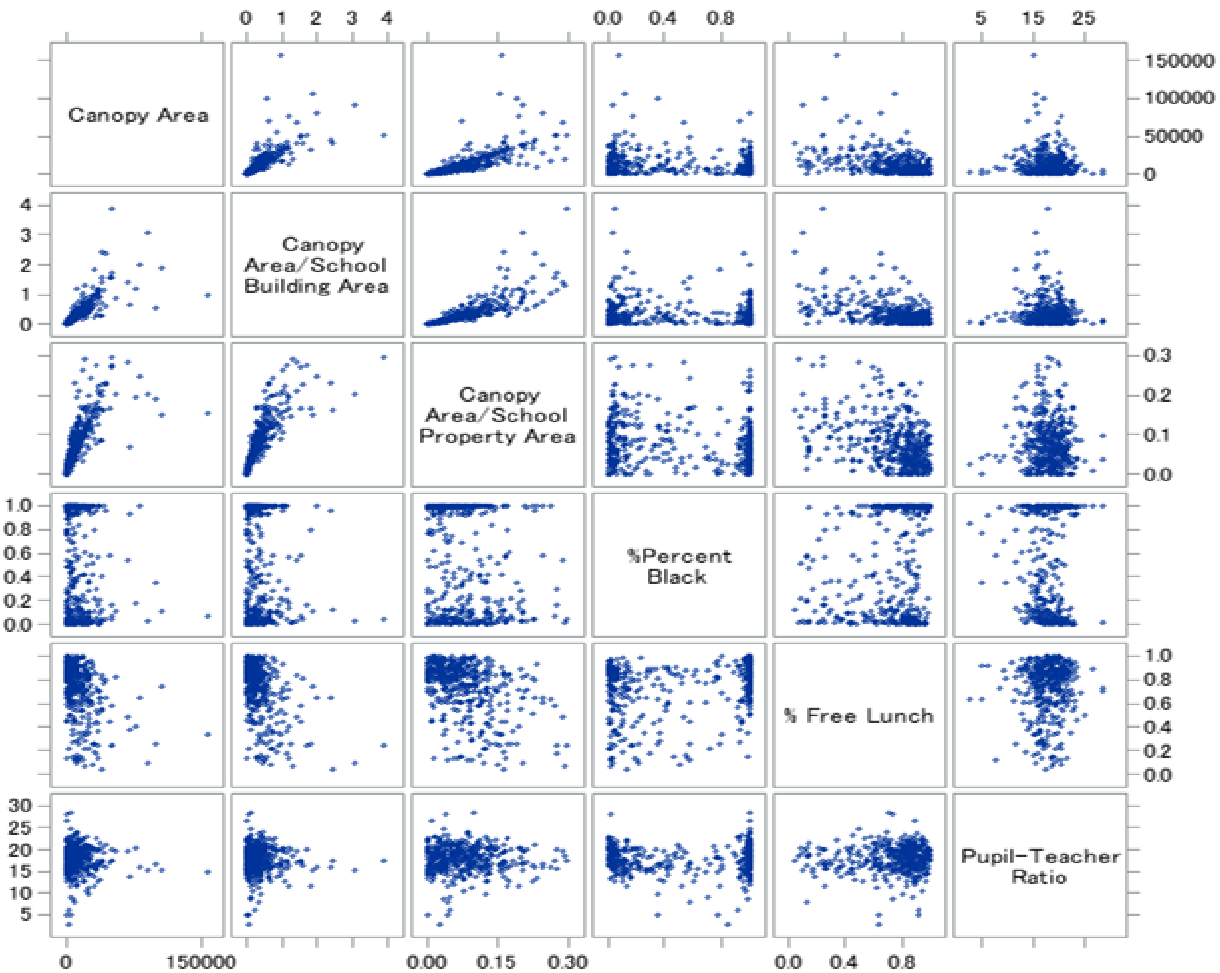


Figure 12. Scatter plot for school-level variables

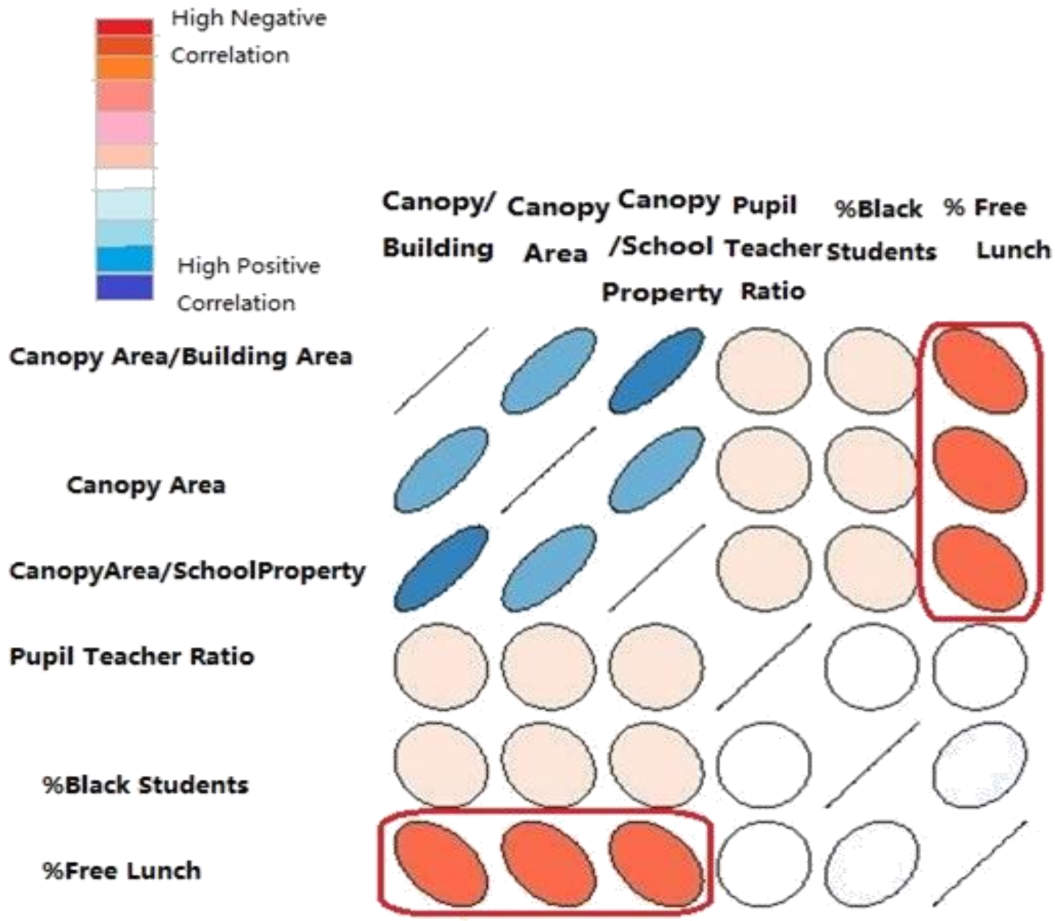


Figure 13 Correlation heat map for school-level variables

Clearly, there are strong positive correlations between all the three measurements of greenness, which suggests only one measurement at a time should be included in the model. Meanwhile, the strong negative correlation between percent of free lunch and all the three measurements of greenness indicates the existence of collinearity problem. On one hand, because the effect of greenness is central in to the analysis, we don't want any other variable to confound with variables related with greenness. On the other hand, we don't want to lose the whole information of free lunch by excluding the variable Percent of Free Lunch (Pfree) totally. Instead, we utilize this variable as the criteria to stratify the 420 schools into so-called "Poor" (Percent of Free Lunch>=85%) and "Rich" schools (Percent of Free Lunch<85%). The table below shows the descriptive statistics of school-level variables for these two school categories:

Table 4. School-level descriptive statistics by poor and rich schools

(Poor Schools) Variables	Mean	Std Dev
% Free Lunch	0.9	0.053
Pupil-Teacher Ratio	17.999	2.945
%Percent Black	0.63	0.433
Canopy Area	8088.299	8098.657
Canopy Area/School Property Area	0.058	0.046
Canopy Area/School Building Area	0.215	0.209

(Rich Schools) Variables	Mean	Std Dev
% Free Lunch	0.581	0.209
Pupil-Teacher Ratio	17.849	3.243
%Percent Black	0.482	0.417
Canopy Area	17945.53	20046.605
Canopy Area/School Property Area	0.095	0.066
Canopy Area/School Building Area	0.484	0.527

From the table above we can see significant gaps between the means of all measurements of greenness. All measurements for rich schools are almost double times of those for poor schools. Also, the variances of all the measurements for rich schools are much larger than those of poor schools. The figure below shows the different distributions of these measurements between poor (upper part) and rich schools (lower part). Therefore, it makes sense to separate the schools into these two categories. In our later analysis, we will always treat poor schools' data and rich schools' data as separate.

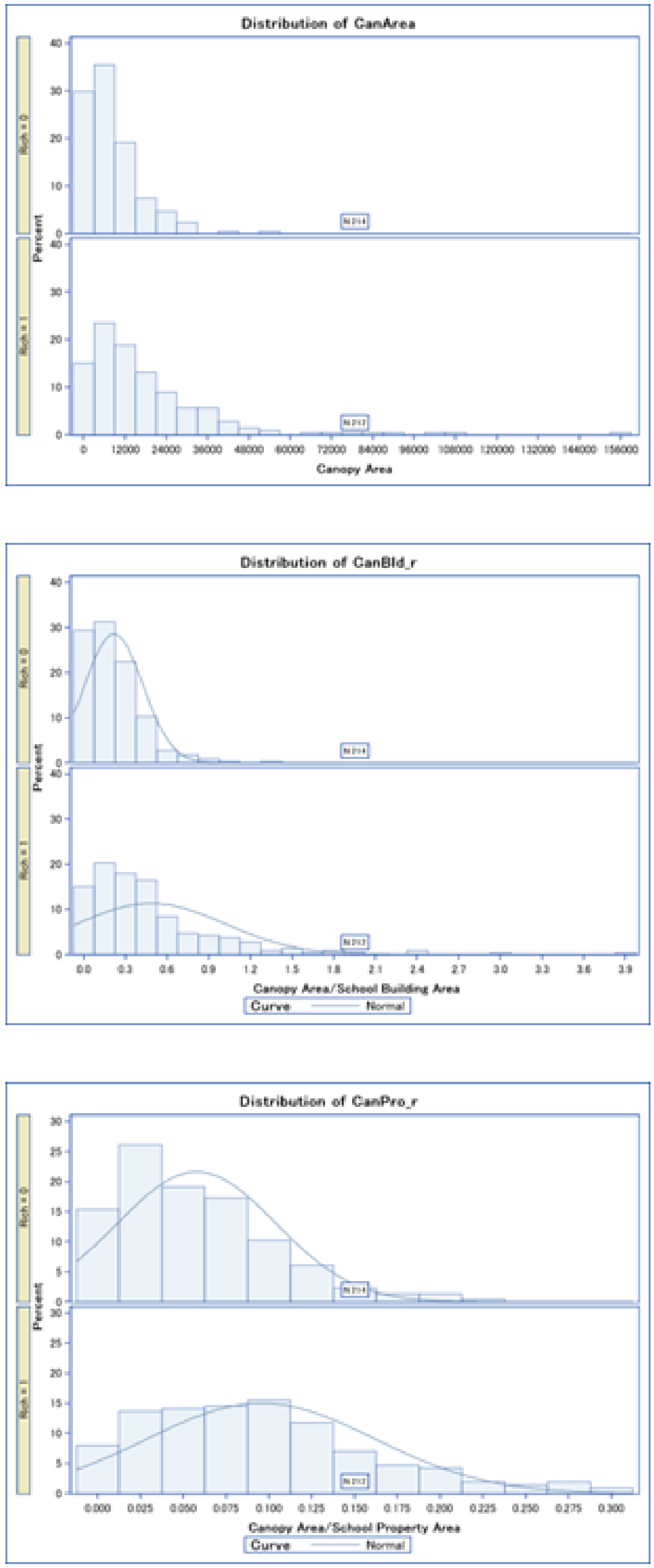


Figure 14. Comparison of histograms of CanArea, CanBld_r, CanPro_r between poor (upper) and rich (lower) schools

3.2 Student-Level Data

Student-level variables are summarized in Table 5. The variables include all possible predictors that affect a student's performance in school, and the independent variable ITBS Reading Scores as a measure for a student's performance in school in our model. For the 420 schools, there are totally 173842 students.

Table 5. Student-level variables

Variable Name	Description	Values
schno	School ID	Numeric
ITBSRDSS	ITBS Reading Scores	Numeric
Grade	Student Grade	Numeric (3-8)
C3_new	Student age correction factor	1= Older than normal age for their grade 0= Otherwise
Gender	Student gender	0= Male 1= Female
Lunch	Free lunch status	0=No free lunch 1=Free Lunch
Race	Student Race	Base = Black Dummy 1: Hispanic Dummy 2: White, American, Indian or Asian
SE	Special Education Status	0=Not in special education 1=All other special education groups
BC	Student Bilingual Category	0=Not in Bilingual program 1=Bilingual program

We still separate the student-level data as from poor schools and from rich schools. The pie charts and bar charts below show the composition of students in poor and rich schools.

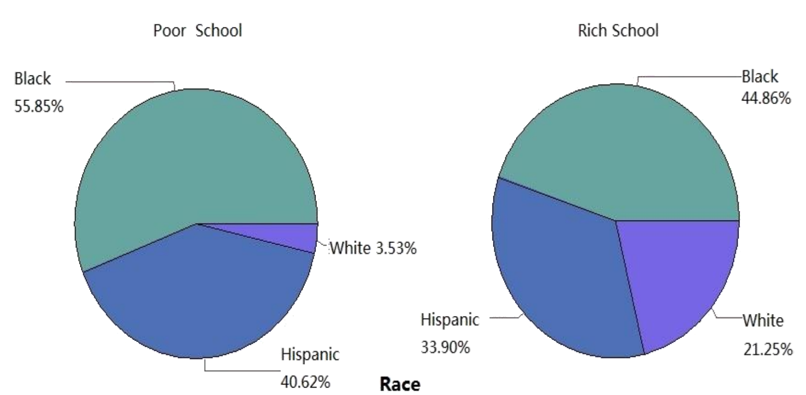


Figure 15. Race composition of poor schools (left) and rich schools (right)

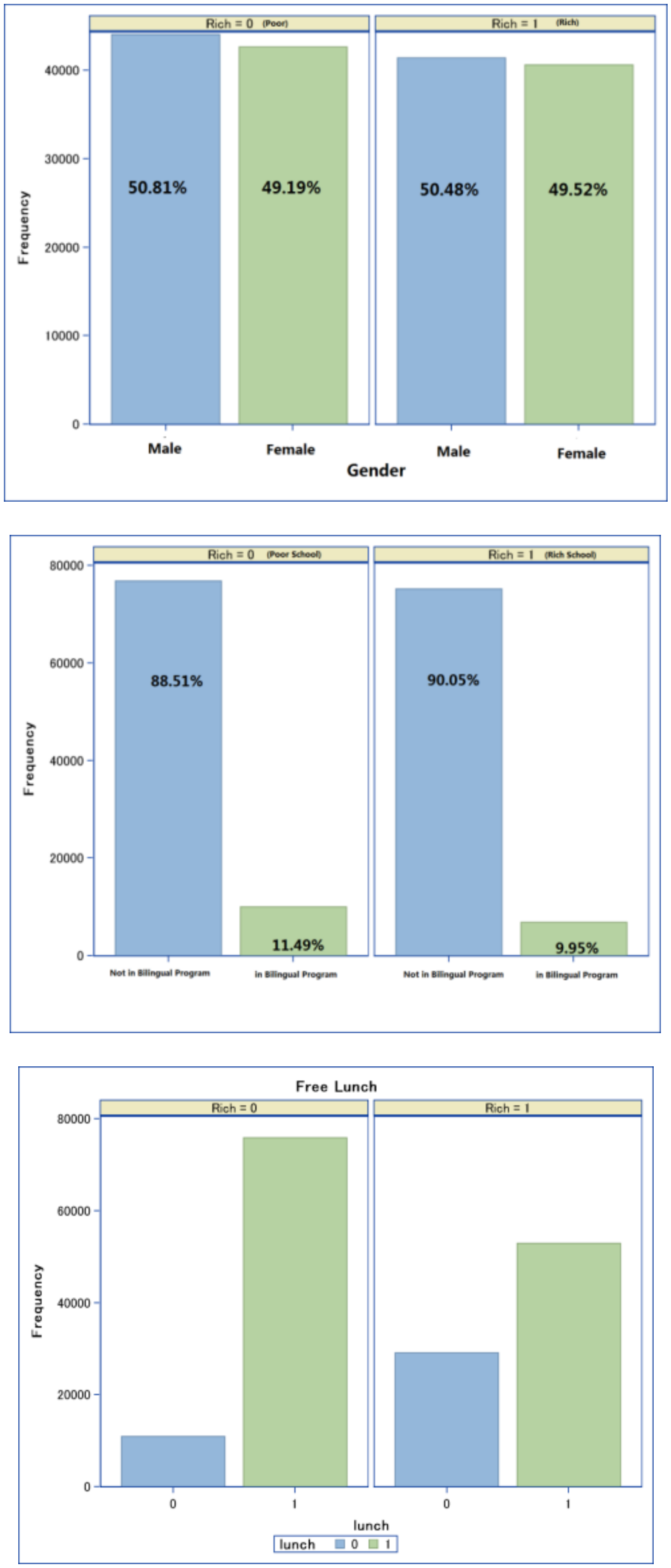


Figure 16. Gender, bilingual and free lunch by poor and rich schools

For most of the variables, there are no big difference between the composition of students from rich schools and poor schools. The two variables that have big differences are Race and Free Lunch Status. We can see from figure 15, the percent of white students are quite small. Thanks to a very large dataset, there are still total 3062 white students in poor schools. So we don't need to worry about sample size problem.

For the ITBS reading scores, we can see the distributions showed in the right figure are quite normal for both poor schools and rich schools, but rich schools' students have a much better performance in average. The normality of the independent variable makes our model more reliable.

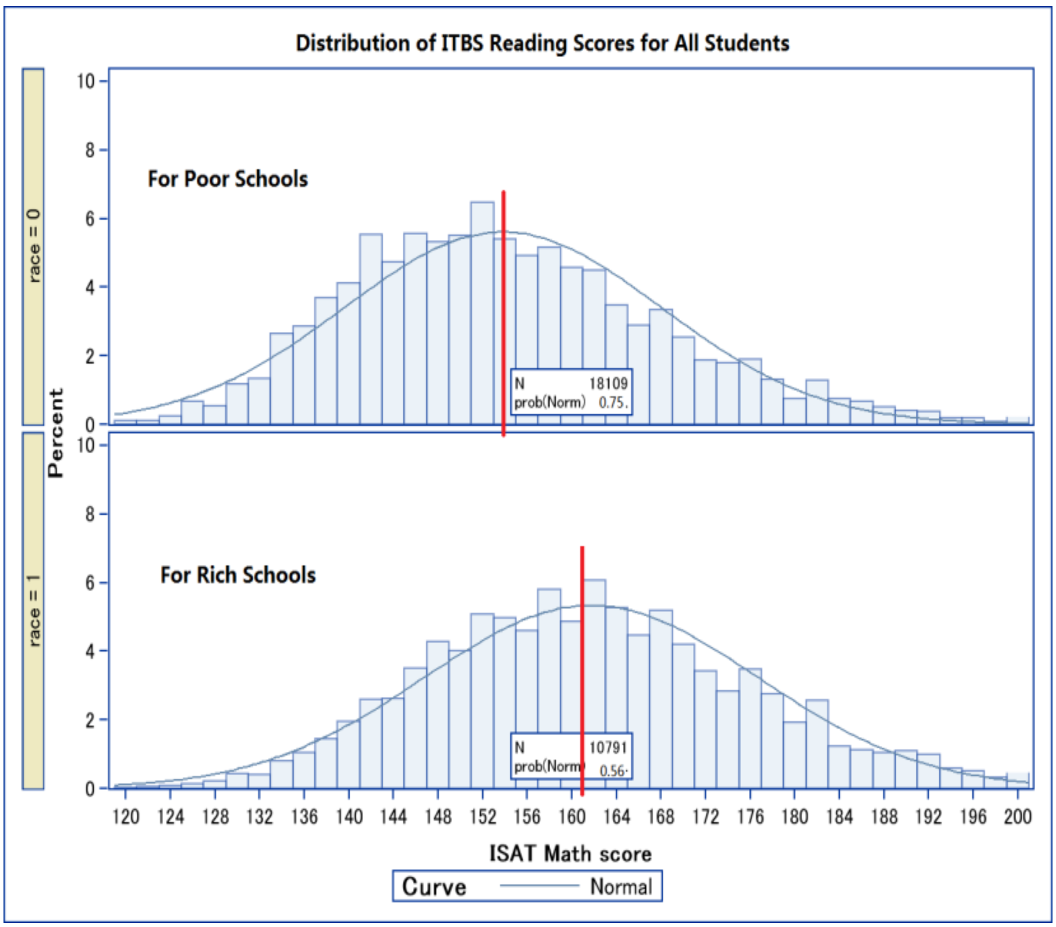


Figure 17. Distribution of ITBS Reading Scores between poor schools (upper) and rich schools (lower)

Similarly, we also need to check the collinearity problem between the student-level variables. The correlation heat map below shows the strong correlation between Age and Grade.

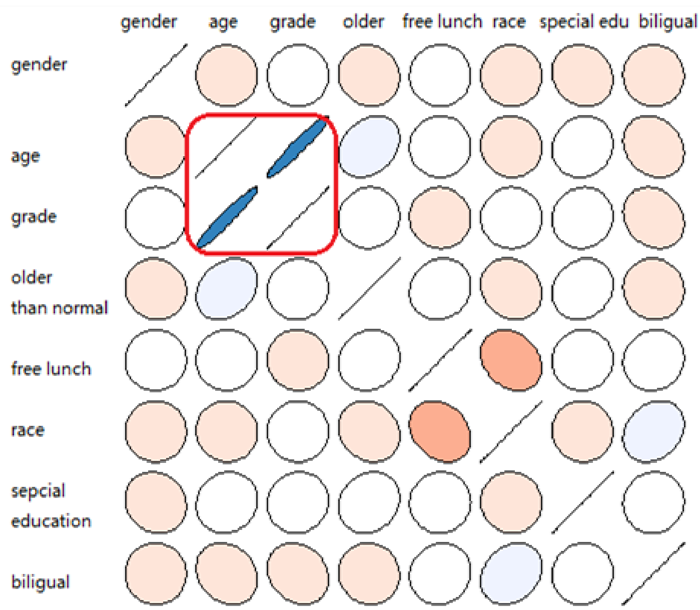


Figure 18. Student-level variables correlation heat map

We run the ordinary least squares regression with ITBS Reading Score as independent variable and all other student-level variables as predictors (This model is essentially the level-1 model in Hierarchical Linear Model) and check the VIF(variance inflation factor quantifies the severity of collinearity in an ordinary least squares regression analysis). The extreme large VIFs for Age and Grade in table below also show the redundancy of including both Age and Grade:

Table 6. VIF analysis

Parameter Estimates						
Variable	**DF**	**Parameter Estimate**	**Standard Error**	**t Value**	**Pr > |t|**	**Variance Inflation**
Gender	1	2.13093	0.11913	17.89	<.0001	1.02022
Age	1	-0.78567	0.18375	-4.28	<.0001	**31.28774**
Grade	1	13.11748	0.18636	70.39	<.0001	**28.89403**
C3_new	1	-11.35608	0.24153	-47.02	<.0001	2.74423
Free Lunch	1	-10.87898	0.14535	-74.85	<.0001	1.09840
Race	1	9.60141	0.09179	104.61	<.0001	1.16335
Special Education	1	-26.53538	0.17656	-150.29	<.0001	1.04722
Bilingual	1	-19.91228	0.21085	-94.44	<.0001	1.13357

After excluding Age and running the ordinary least square model again, all the VIFs are around 1 which suggests no severe collinearity problem in student-level variables. Both the school-level data and student-level data are now ready to use to build the Hierarchical Linear Model.

4. HIERARCHICAL LINEAR MODEL

4.1 Introduction

HLM is a statistical model that allows to specify and to estimate the relationship between variables. We use HLM when we observe a hierarchical data structure, which means the data is clustered or nested. Figure 19 illustrates a typical HLM example from a 3-level data structure.

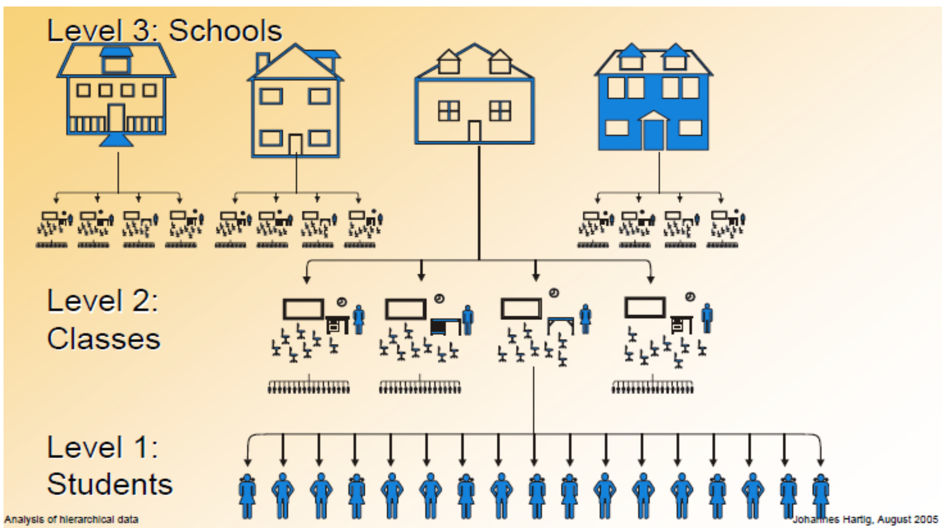


Figure 19. 3-level data structure

Suppose we collect information from students, classes and schools, and we want to conduct some research on the student achievement. Because some students come from the same class, they share some common factors, such as the same instructor and class environment. However, these common factors may be different from class to class. Therefore, the student data may violate independence assumption of OLS, which requires each student is independent to each other. This is the reason why we use HLM, which take these effects into account.

4.2 Model Settings

In this project, the dataset contains two levels, level 1 (student level) and level 2 (school level). To build a two-level HLM, let's look at the level-1 model first.

$$Level\ 1: Y = \beta_0 + \beta_1 X_1 + \cdots + \beta_p X_p + \varepsilon \quad (student\ level, random\ effect)$$

In level-1 model, Y is the response from student we are interested in, which in this project is the student's testing score. For the X's, we use student-level variables including Gender, Grade, Bilingual Category etc. as our predictors. Here, the β's are the effects of predictors. The main difference between a hierarchical linear model and OLS is that these β's are now random effects. It means they may be different for students from different schools and this is why we need the level-2 models, which to model the random effects in level 1. Next, let's look at level-2 model.

$Level\ 2$

$$\beta_0 = \gamma_{00} + \gamma_{01} Z_1 + \cdots + \gamma_{0p} Z_p + \mu_0 \quad (school\ level, fixed\ effect)$$

$$\beta_1 = \gamma_{10} + \gamma_{11} Z_1 + \cdots + \gamma_{1p} Z_p + \mu_1$$

$$\vdots$$

$$\beta_p = \gamma_{p0} + \gamma_{p1} Z_1 + \cdots + \gamma_{pp} Z_p + \mu_p$$

In level 2, we model the random effects in level 1 (the $\beta's$), using school-level variables(the Z's) including percent of black student, percent of free lunch and one measurement of greenness. Here, the $\gamma's$ are called fixed effect because it is the highest level, and no more randomness from higher level.

4.3 Model Assumptions

Before we conduct the analysis, let's compare the differences in basic assumptions between HLM and OLS. First, for OLS, the function must be in linear form. Also, the residuals should follow normal distribution with constant variance and observations are independent to each other. On the other hand, for HLM, the function forms are linear at each level; level-1 residuals are normally distributed and level-2 residuals (the u's) follow multivariate normal distribution; constant variance for level-1 residuals; level-1 residuals and level-2 residuals are uncorrelated and observations at highest level are independent. We should check those assumptions on HLM when we conduct the analysis to justify our result.

4.4 Predictors Review

Table 2 and Table 5 list the predictors for level-1 and level-2 models respectively. In our analysis, we use percent of free lunch to separate the schools into rich and poor schools which is consistent with the previous study. Here, the cutoff point is 85%, which is the medium of variable Pfree. Also we centered the level-2 (school-level) variables by the grand mean out of interpretation purpose. Centering is not applicable for level-1 (student-level) variables because most of them are categorical variables. Since we have two groups of schools, poor and rich, and three GIS based greenness measurements, we have six combinations in total. The analysis below used data for poor schools and Canopy Area as greenness measurement.

Using the predictors in both levels, the hierarchical linear model is:

$level\ 1$

$$Y = B0 + B1*(GRADE) + B2*(LUNCH) + B3*(SE) + B4*(GENDER) + B5*(BC) + B6*(C3\_NEW) + B7*(WHITE) + B8*(HISP) + R$$

$level\ 2$

$$B0 = G00 + G01*(PUPTCH) + G02*(PBLACK) + G03*(CANAREA) + U0$$

$$B1 = G10 + G11*(PUPTCH) + G12*(PBLACK) + G13*(CANAREA) + U1$$

$$\vdots$$

$$B7 = G70 + G71*(PUPTCH) + G72*(PBLACK) + G73*(CANAREA) + U7$$

$$B8 = G80 + G81*(PUPTCH) + G82*(PBLACK) + G83*(CANAREA) + U8$$

4.5 HLM report (Poor Schools & Canopy Area example)

Again, we use model based on poor schools and Canopy Area data to illustrate how to read the report. We first look at the level-1 report, which indicates the variance component analysis. In level-1 report, we do not estimate the $\beta's$ since they are random. Instead, we check the assumption of randomness. Table 7 shows the χ^2 test statistic and p-value. As we can see from the table, all p-values are less than .05. So we reject the null hypothesis that β_i is not random which means the randomness assumption on level-1 model predictors is valid. In our analysis, the assumption is adequate for all the models based on six combinations of data.

Table 7. Variance component table

Random Effect	Std	Var	df	Chi-sq	P-value
INTRCPT1, U0	7.12	50.70	99	233.28	0.000
GRADE slope, U1	1.24	1.56	99	455.37	0.000
LUNCH slope, U2	2.65	7.06	99	162.05	0.000
SE slope, U3	5.97	35.70	99	281.21	0.000
GENDER slope, U4	1.13	1.27	99	132.02	0.015
BC slope, U5	2.54	6.47	99	169.11	0.000
C3_NEW slope, U6	2.17	4.74	99	163.01	0.000
WHITE slope, U7	4.24	18.02	99	149.27	0.001
HISP slope, U8	3.47	12.10	99	173.54	0.000
level-1, R	22.52	507.36			

Next, we look at the report for level-2 model. Table 8 is a partial output for the following level-2 equations.

$$B0 = G00 + G01 * (PUPTCH) + G02 * (PBLACK) + G03 * (CANAREA) + U0$$

$$B1 = G10 + G11 * (PUPTCH) + G12 * (PBLACK) + G13 * (CANAREA) + U1$$

$G00, G01, G02, and\ G03$ are related to B0, which is the intercept of level-1 model. G00 shows on average B0 equals to 148.59, given PUPTCH, PBLACK, and CANAREA at their grand means. This number of 148.59 makes sense because it shows the testing score of a baseline student in an average school. G01 shows that, given PBLACK and CANAREA fixed, one unit increase in PUPTCH will decrease B0 by .51 units. Similarly, G02 shows that, given PUPTCH and CANAREA fixed, one unit increase in PBLACK will decrease B0 by 7.62 units. G03 shows that, given PUPTCH and PBLACK fixed, one unit increase in CANAREA will increase B0 by .000027 units. Notice, the unit of canopy area is square feet and that is why G03 is such a tiny number. Once it multiplies with the canopy area, which is a large number, the effect is considerable. Next, let's look at G10, G11, G12, and G13. G10 indicates that in average B1, the slope of GRADE, equals to 11.83, given PUPTCH, PBLACK, and CANAREA are at their grand means. The number 11.83 makes sense because it shows that when a baseline student goes from third grade to fourth grade, the testing scores will increase 11.83, in average. G11 shows that, given PBLACK and CANAREA fixed, one unit increase in PUPTCH will increase B1 by .04 units. Similarly, G12 shows that, given PUPTCH and CANAREA fixed, one unit increase in PBLACK will increase B1 by .49 units. G13 shows that, given PUPTCH and PBLACK fixed, one unit increase in CANAREA will increase B1 by .000017 units.

After we know how to interpret the coefficients, let's look at the p-values. If we use $\alpha = .05$, among G00, G01, G02, and G03, only G03 is insignificant, which shows the greenness effect on B0, the intercept of level-1 model is not strong. On the other hand, among G10, G11, G12, and G13, G11 and G13 are insignificant, which indicates that pupil-teacher ratio effect and greenness effect are not strong for GRADE slope in level-1 model.

Table 8. Partial fixed effect table

Fixed Effect	Coef.	std	T-ratio	df	P-value
For INTRCPT1, B0					
INTRCPT2, G00	148.59	0.68	218.21	205	0.000
PUPTCH, G01	-0.51	0.25	-1.99	205	0.047
PBLACK, G02	-7.62	1.50	-5.05	205	0.000
CANAREA, G03	0.000027	0.000084	0.31	205	0.750
For GRADE slope, B1					
INTRCPT2, G10	11.83	0.10	112.90	205	0.000
PUPTCH, G11	0.04	0.03	1.12	205	0.261
PBLACK, G12	0.49	0.24	2.04	205	0.042
CANAREA, G13	0.000017	0.000013	1.30	205	0.194

Now, since the main concern of this project is the greenness effect. We summarize the greenness effect in this model in the following Table 9. From the p-values listed below, we find the greenness effects are not significant in the poor schools & canopy area model. It is only on the margin for the WHITE slope, which compares the difference between White and Black students. It shows there is slightly evidence that the difference between White and Black is getting larger in a greener school.

Table 9. Poor & CanArea (= Canopy Area)

Fixed Effect	Coef.	std	t	df	P-value
For INTRCPT1, B0	0.000027	0.000084	0.318	205	0.750
For GRADE slope, B1	0.000017	0.000013	1.303	205	0.194
For LUNCH slope, B2	-0.000007	0.000039	-0.169	205	0.866
For SE slope, B3	-0.000045	0.000062	-0.727	205	0.468
For GENDER slope, B4	0.000004	0.000022	0.194	205	0.847
For BC slope, B5	-0.000041	0.000048	-0.863	205	0.389
For C3_NEW slope, B6	-0.000022	0.000041	-0.548	205	0.584
For WHITE slope, B7	0.000114	0.000062	1.846	205	0.066
For HISP slope, B8	0.000011	0.000038	0.297	205	0.767

4.6 Analysis of greenness effect

4.6.1 Poor schools

Now let's look at the greenness effects using different combinations of data. First we show the models based on data for poor schools and the other two greenness measurements, which are canopy area divided by school property area and canopy area divided by building area.

Table 10. Poor & CanPro_r (= Canopy Area / School Property Area)

Fixed Effect	Coef.	std	t	df	P-value
For INTRCPT1, B0	9.08	14.18	0.640	205	0.522
For GRADE slope, B1	3.25	2.24	1.445	205	0.150
For LUNCH slope, B2	-5.94	6.67	-0.891	205	0.374
For SE slope, B3	0.97	10.53	0.092	205	0.927
For GENDER slope, B4	3.89	3.85	1.010	205	0.314
For BC slope, B5	-12.36	7.53	-1.640	205	0.102
For C3_NEW slope, B6	-0.54	5.43	-0.100	205	0.921
For WHITE slope, B7	13.92	15.26	0.912	205	0.363
For HISP slope, B8	-3.09	8.97	-0.346	205	0.730

Table 11. Poor & CanBld_r (= Canopy Area / School Building Area)

Fixed Effect	Coef.	Std	t	df	P-value
For INTRCPT1, B0	2.570126	3.554679	0.723	205	0.470
For GRADE slope, B1	0.683227	0.594935	1.148	205	0.253
For LUNCH slope, B2	-0.706640	1.646673	-0.429	205	0.668
For SE slope, B3	-1.117042	2.180709	-0.512	205	0.609
For GENDER slope, B4	0.437931	0.851087	0.515	205	0.607
For BC slope, B5	-3.267612	1.983633	-1.647	205	0.101
For C3_NEW slope, B6	-0.764488	1.346956	-0.568	205	0.571
For WHITE slope, B7	3.121354	2.808051	1.112	205	0.268
For HISP slope, B8	-1.107233	1.748019	-0.633	205	0.527

From Table 10 and Table 11 above, if we chose 0.05 significance level, none of the greenness effect is significant, which indicates that the greenness has no apparent effect on student's achievement in poor schools.

4.6.2 Rich schools

Now, let's look at the results from rich schools & three greenness measurements.

Table 12. Rich & CanArea (= Canopy Area)

Fixed Effect	Coef.	Std	t	df	P-value
For INTRCPT1, B0	-0.000017	0.000043	-0.390	207	0.697
For GRADE slope, B1	0.000026	0.000006	4.130	207	0.000
For LUNCH slope, B2	-0.000023	0.000014	-1.622	207	0.106
For SE slope, B3	-0.000046	0.000025	-1.836	207	0.067
For GENDER slope, B4	0.000027	0.000013	2.143	207	0.033
For BC slope, B5	-0.000027	0.000019	-1.396	207	0.164
For C3_NEW slope, B6	0.000040	0.000017	2.288	207	0.023
For WHITE slope, B7	-0.000007	0.000029	-0.246	207	0.806
For HISP slope, B8	0.000022	0.000021	1.074	207	0.284

Table 13. Rich & CanPro_r (= Canopy Area / School Property Area)

Fixed Effect	Coef.	Std	t	df	P-value
For INTRCPT1, B0	-14.131492	10.227457	-1.382	207	0.169
For GRADE slope, B1	8.688403	1.723554	5.041	207	0.000
For LUNCH slope, B2	-12.617955	3.680582	-3.428	207	0.001
For SE slope, B3	-11.044282	6.115842	-1.806	207	0.072
For GENDER slope, B4	10.960424	2.907817	3.769	207	0.000
For BC slope, B5	0.624030	6.799259	0.092	207	0.927
For C3_NEW slope, B6	6.982204	4.313888	1.619	207	0.107
For WHITE slope, B7	2.333384	7.790031	0.300	207	0.765
For HISP slope, B8	6.379551	6.453419	0.989	207	0.324

Table 14. Rich & CanBld_r (= Canopy Area / School Building Area)

Fixed Effect	Coef.	Std	t	df	P-value
For INTRCPT1, B0	-1.322723	1.254208	-1.055	207	0.293
For GRADE slope, B1	0.951970	0.209186	4.551	207	0.000
For LUNCH slope, B2	-1.312652	0.451811	-2.905	207	0.005
For SE slope, B3	-1.239337	0.740097	-1.675	207	0.095
For GENDER slope, B4	1.294681	0.343201	3.772	207	0.000
For BC slope, B5	-1.448196	0.828751	-1.747	207	0.082
For C3_NEW slope, B6	1.539419	0.606118	2.540	207	0.012
For WHITE slope, B7	0.831659	0.943588	0.881	207	0.379
For HISP slope, B8	1.124652	0.945989	1.189	207	0.236

From Table 12, Table 13 and Table 14, if we still use significance level of 0.05, we see a different story in rich schools. First, we notice the greenness effects on GRADE, SE, and GENDER are significant using all three greenness measurements. It shows that greenness does affect a student's achievement in rich schools via these three predictors. First, for GRADE, the coefficients are all positive in three models. It shows that when students go to higher grade in a greener school, the benefit of greenness is getting larger. Next, for SE, it represents whether a student takes special education. We see from the table that the average SE effects are negative and the greenness effects are also negative in all three models. It indicates that a student who takes special education has lower ITBS reading score than a student who does not take special education. Moreover, the difference is getting larger in a greener rich school. Last, for GENDER, the positive number shows in average, a female student has higher ITBS reading score than a male student, and the positive greenness effects in all three models show that this difference is getting larger in a greener rich school.

For GRADE, SE, and GENDER, the greenness effects agree with each other using all three measurements. There are some greenness effects that are significant under some greenness measurement but not in the others. For example, the greenness effect on LUNCH is significant when we use canopy divided by school property area and canopy divided by building area, but it is insignificant when we use canopy area as greenness measurement. In general, we can say the greenness effect is significant on GRADE, LUNCH, SE, GENDER, BC, and C3_NEW in at least in one greenness measurement, but not on the intercept and on the race slopes.

4.7 Comparison

4.7.1 Poor versus rich schools

From the analysis of poor and rich schools, we find some commons and differences. First, let's look at the commons. We can see from the tables in the previous section, the greenness effects are not significant on the intercepts. This indicates that we have no evidence to say the greenness affects the average of students' achievement directly. Also, all greenness effects on race slopes are insignificant at $\alpha = .05$. It shows that there is no evidence to say the difference between White and Black student or Hispanic and Black student are getting larger in a greener school.

On the other hand, let's look at the differences. The greenness effects are not significant in poor schools if we use significance level 0.05. It shows that a greener poor school will not help the students' achievement with respect to other poor schools. However, the greenness effects are significant in the rich schools models on many of the level-1 predictors. It shows that a greener rich school does help the students' achievement.

4.7.2 GIS based analysis versus subjective rating analysis

By comparing our findings with the previous study, we also find some commons and differences. The main similarity of results is that the greenness effect is important among rich schools but not among poor schools, i.e. greenness helps students' achievement in rich schools but not in poor schools. On the other hand, for the difference between GIS based versus subjective rating analysis, we find some cases where greenness effects have opposite signs. For example, greenness effect on BC is positive in the previous study but negative in the GIS based study. We do expect some discrepancy between two studies because we use totally different greenness measurements.

4.8 Diagnostic Analysis

From HLM7 manual p.35, we know that the HLM report contains two tables for the final estimation of fixed effects. One provides the model-based estimates of standard errors, and the other provides robust estimates of stand errors. If the robust and model-based standard errors are substantively different, it is recommended that the tenability of key assumptions should be investigated further. For all six models we build in this project, they all pass this criterion. Therefore we do not worry about violation of the basic assumptions. Meanwhile, because we have level-1 and level-2 residuals, we also display some graphics to check the basic assumptions. Notice, it is just a side justification, and based on the manual, we should not have to worry about the validation of our models. The Hierarchical Linear Model we use is a two-level model. The level-1 model residuals are ITBS Reading Score residuals. There are 8 packs of level-2 model residuals since we have 8 level-2 models regarding with the 8 student-level predictors.

The model has several main theoretical assumptions:

- Normality: Level-1 model residuals are normally distributed and Level-2 model residuals follow multivariate normal distribution.
- Homoscedasticity: constant variance for level-1 model residuals within each school.
- Independence: Level-1 model residuals and level-2 residuals are uncorrelated.

The diagnostic analysis focuses on checking whether these assumptions are valid for our model using residual data. The residual data in the report is from the model using poor schools data and Canopy Area as greenness measurement.

4.8.1 Normality of Residuals

First, we check the normality of residuals also known as ITBS Reading Score residuals from level-1 model, and Figure 20 shows good normality:

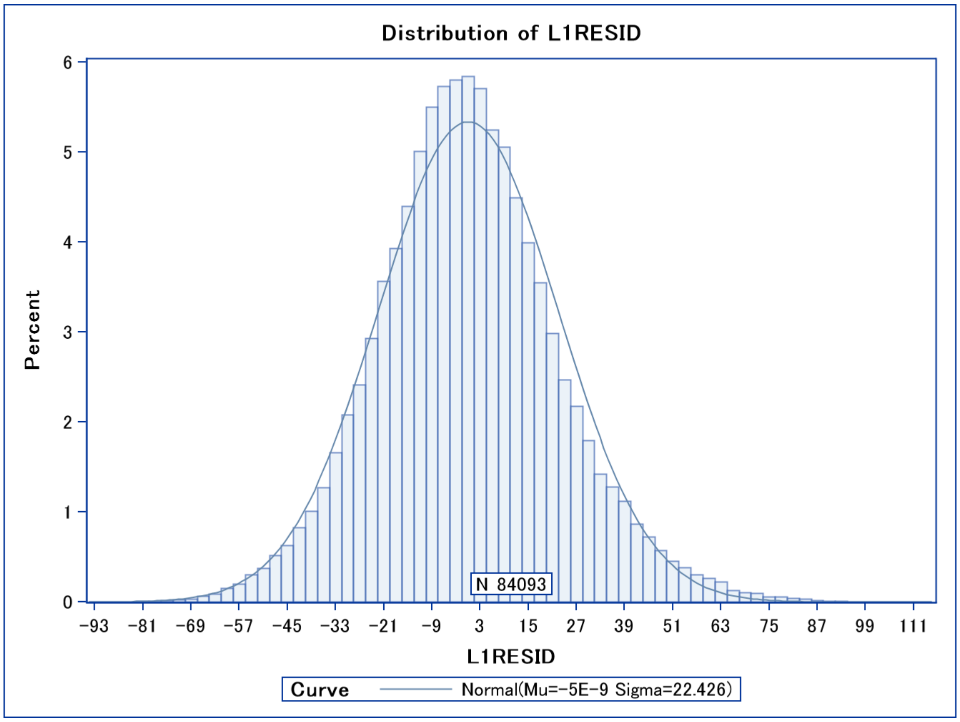


Figure 20. Distribution of student-level residuals

For the multivariate normal distribution assumption for all the level-2 model residuals, there is no easy way to check. Instead, we only check the normality of all the level2 model residuals separately. Figure 21 shows distributions of 8 packs of residuals for Grade, Free Lunch Status, Special Education Status, Gender, Bilingual Category, Hispanic Race, White Race, and Older than Normal Status. None of them violate the normality assumption at 0.1 significance level.

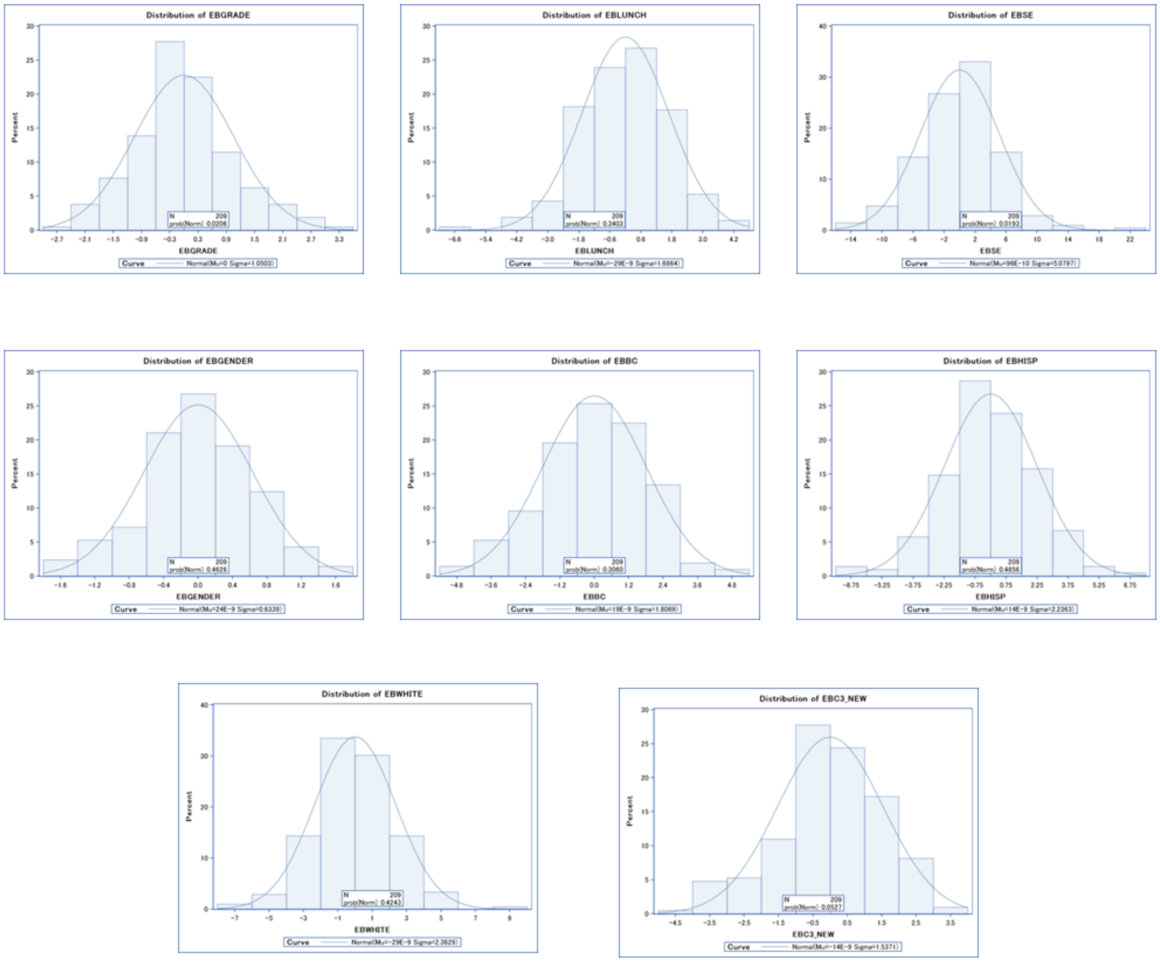


Figure 21. Distribution of school-level residuals

4.8.2 Homoscedasticity

For Homoscedasticity, our assumption states that within each school the residuals of ITBS Reading Scores should relatively have the same variance. Figure 22 below shows the standard deviations of ITBS Score residuals within 209 poor schools. The distribution is quite random and the range of 95% confidence interval is 7.59, which is relatively small comparing with the mean 22.2. So we believe the assumption of constant standard deviations is valid.

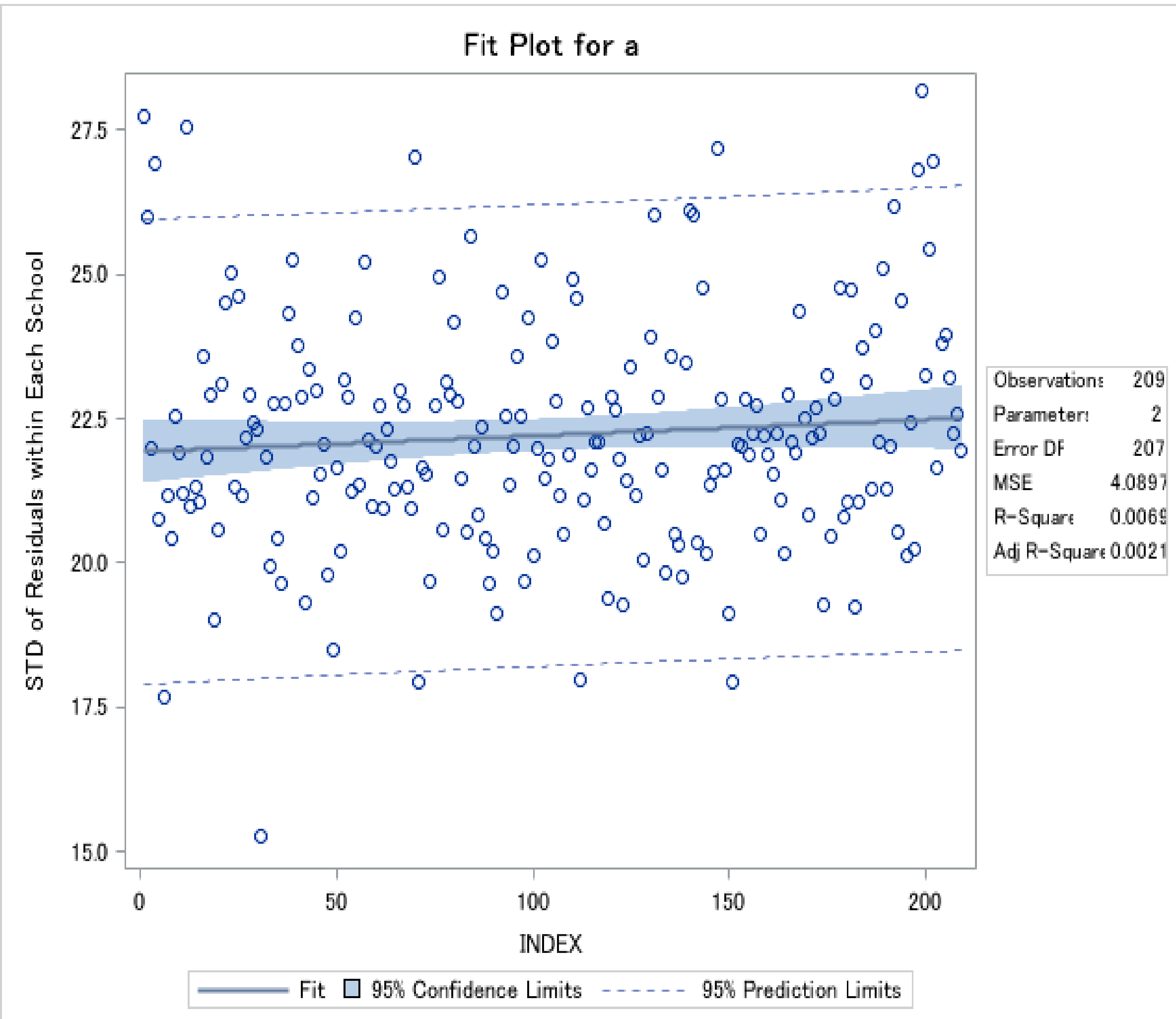


Figure 22. Standard deviations of ITBS Reading Scores in 209 Poor schools

4.8.3 Independence

To check the independence between the level-1 residuals with 8 packs of level-2 residuals, we mainly look at the scatterplots (only 4 of them showed, for Gender, Free Lunch Status, Older than Normal, Bilingual Category).

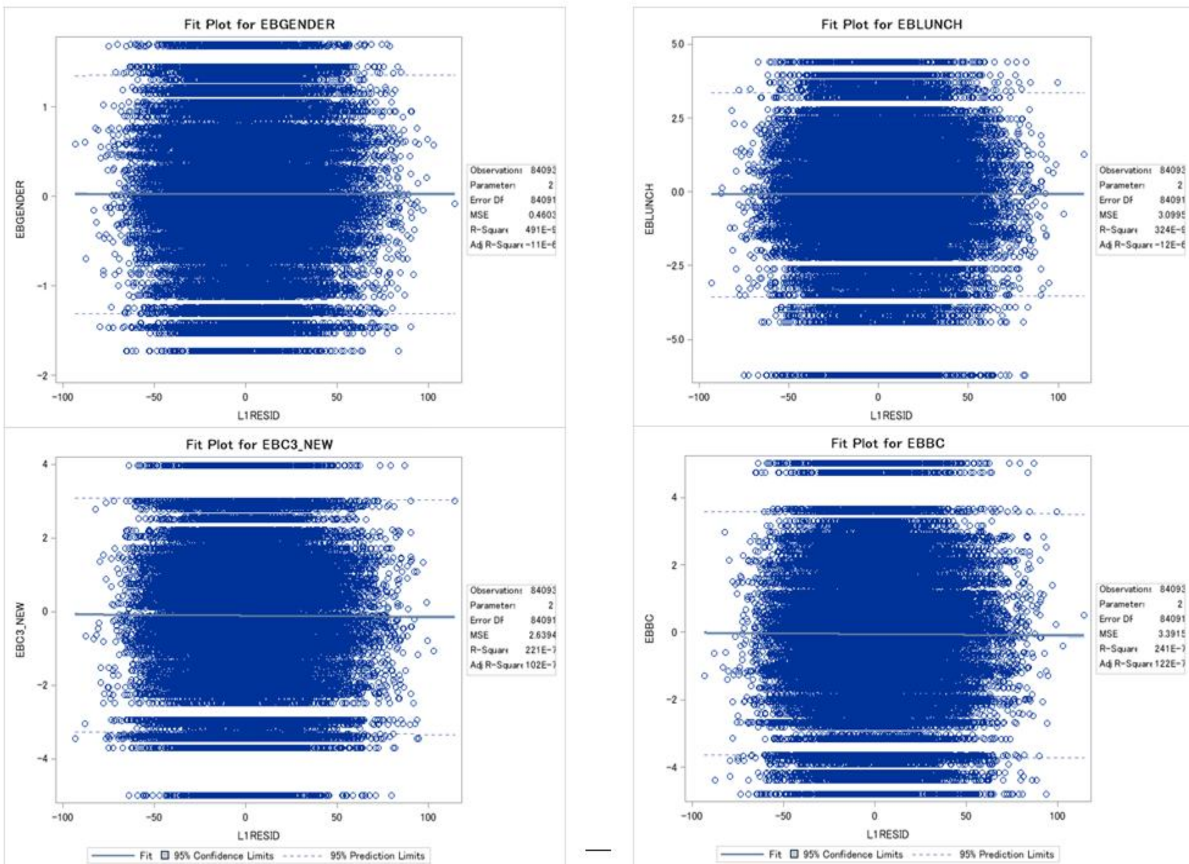


Figure 23. Level-1 residual with level-2 residual

Figure 23 above shows no sign for certain trend or pattern between the level-1 model residuals and level-2 model residuals. The 8 coefficients of correlations between the level-1 model residuals and 8 packs of level-2 model residuals. are all not significant. So we believe this assumption is valid.

For all six models based on six combinations of data, all of them satisfy these three main theoretical assumptions of hierarchal linear model.

5. CONCLUSIONS

5.1 Discussion

Two points became clear from the analysis process. First, the analysis does not include model selection. The reason is that the target of our client is not to find the best model to predict the students' achievement. Instead, the goal is to understand how greenness affects the students' achievement by controlling all relevant predictors. The other point is that there is a time discrepancy issue for our data. In this GIS based analysis, we have the latest data for GIS greenness measurement, which is from 2010-2011. However, we only have the students' testing scores from 2005 as the latest year. The time discrepancy may affect the accuracy of our result. However, how to get the latest data is always a common issue when we conduct research. The main purpose for this project is to help the client to build up a standard operating process for the GIS based study for her future research when she gets the updated data.

5.2 Summary

During this project, we conducted a study on how greenness affects the students' achievement using GIS based greenness measurement and students' testing scores. From our analysis, we find the greenness does help the students' testing scores among rich schools but not among poor schools. In other words, a student has higher testing score in a greener rich school rather than other rich schools; there is no evidence for a student to get higher score in a greener poor school with respect to other poor schools. This finding is consistent to the previous study using subjective greenness measurement. As we discussed in the data preparation section, there are three main limitations of GIS based greenness measurement, the information is restricted inside the school property; it ignores the distance between green coverage and the building; and it does not provide the grass and turf information. However, the value of using GIS based data is that it provides an objective measurement to conduct the research. Using subjective measurement always comes along with human bias and error. Also, it is impossible to have the same person to rate the greenness for a long term research. Therefore, for the future research, we should aim at how to improve the GIS based greenness measurement by considering the advantage of subjective measurement.

REFERENCES

"HLM 7 Hierarchical linear and nonlinear modeling" SSI.

Stephen, W Raudenbush and Anthony S. Bryk "Hierarchical linear models application and data analysis methods second edition"

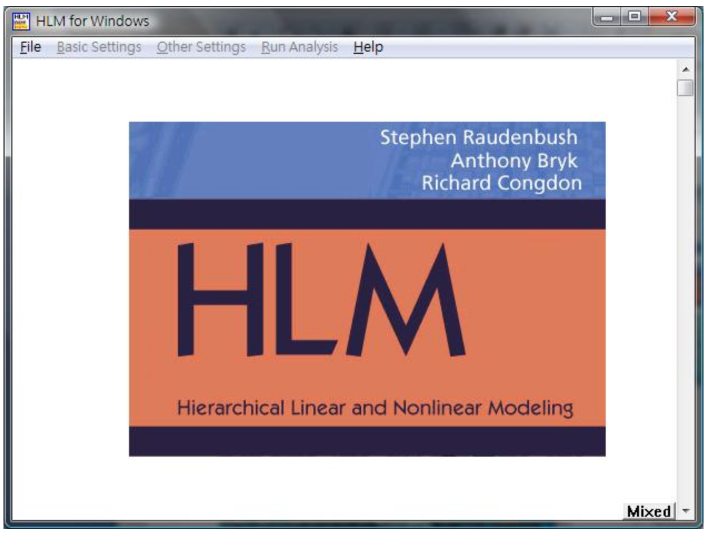
Annie Qu "Comparison of Proc Mixed in SAS and HLM for Hierarchical Linear Models"

APPENDIX

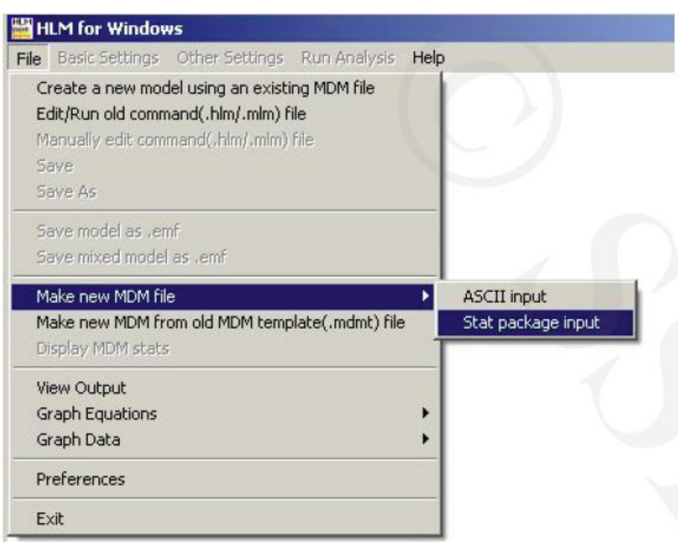
A. HLM Instructions

Two files are needed prior running the HLM analysis: school-level and student-level data. Note that both files need to have the same format (e.g. .sas), length of variables names have to be at most 8 characters, and both datasets need to be sorted in ascending order by school ID.

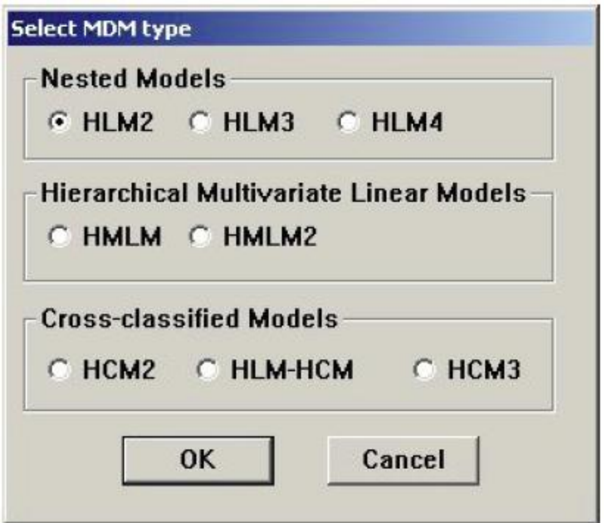
1. Open HLM



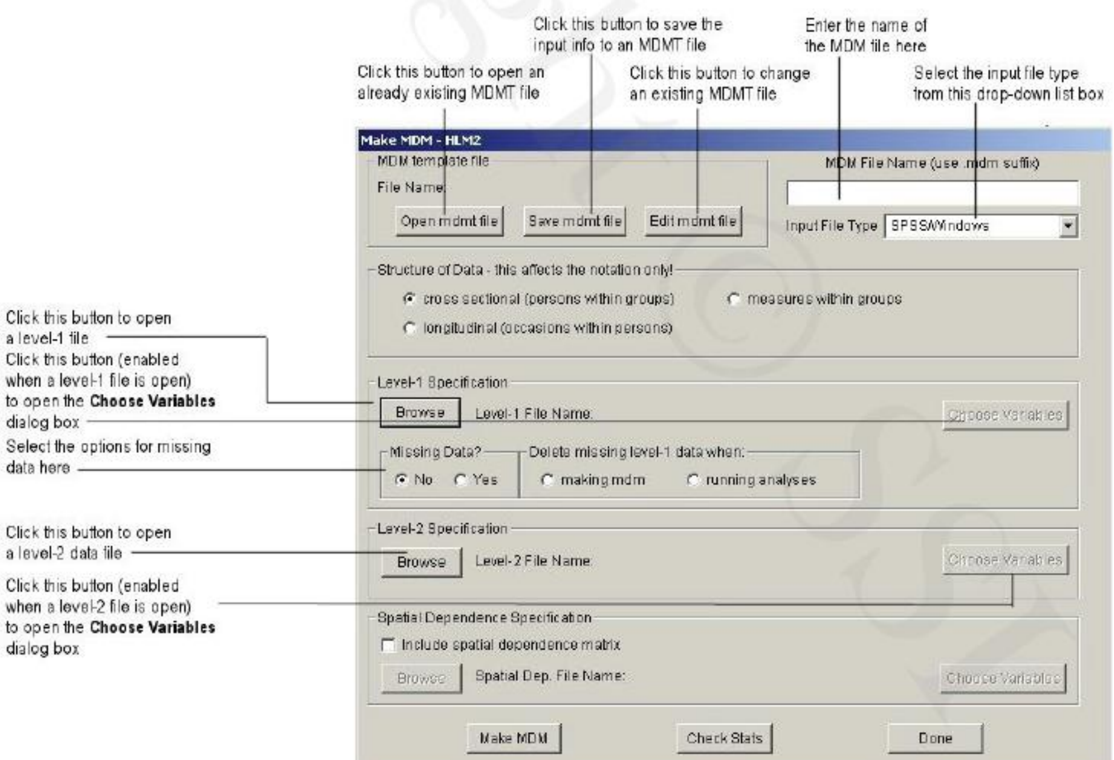
2. Click File → Make new MDM file → Stat Package input



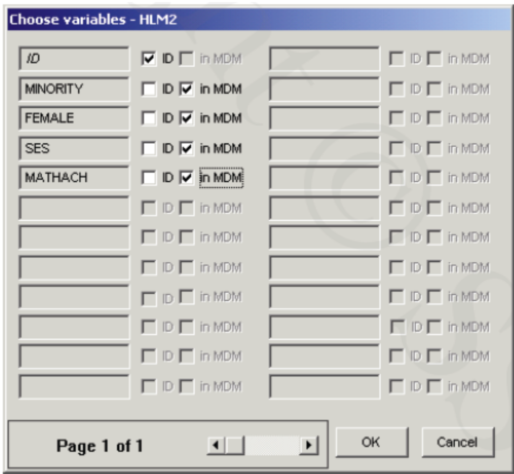
3. Select HLM2 (2 level model)



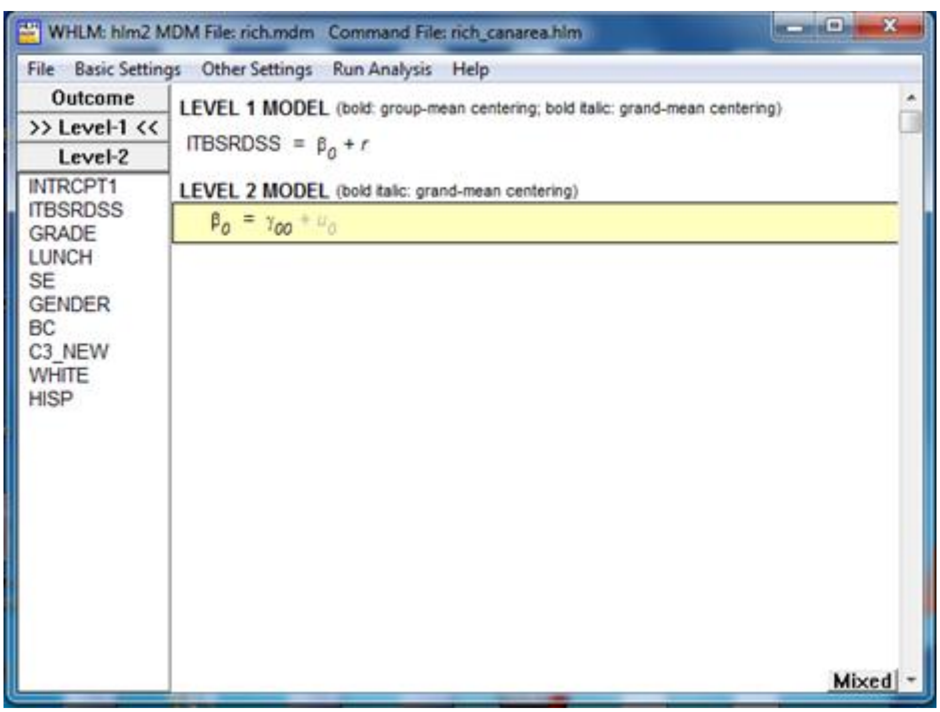
4. Input the data sets. Notice the files should be in the same type. For example, both are SAS files. For missing data, we choose "Yes" and "running analysis." For Input File Type, we choose "Anything else." Save the mdmt file before running the analysis. Also, the files should be under the same folder.



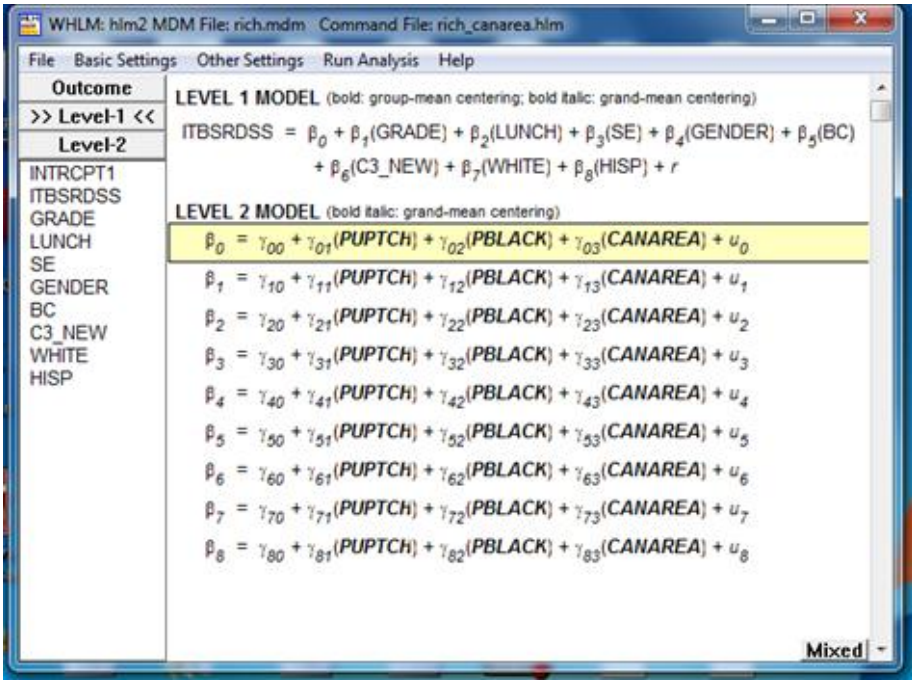
5. Select predictors after input the data. Notice, ID variables should match in level 1 and level 2. It must be ordered. Notice the length of predictors name should be less than 8 characters. Click “Done” after the predictors are selected in both levels.



6. After reading the data, we have to specify our model. First, we specify the level-1 model. By right clicking the predictors on the list, we can specify ITBSRDSS as the response. Next select the predictors in level-1 model. Notice, you can decide to use centered or uncentered predictors.

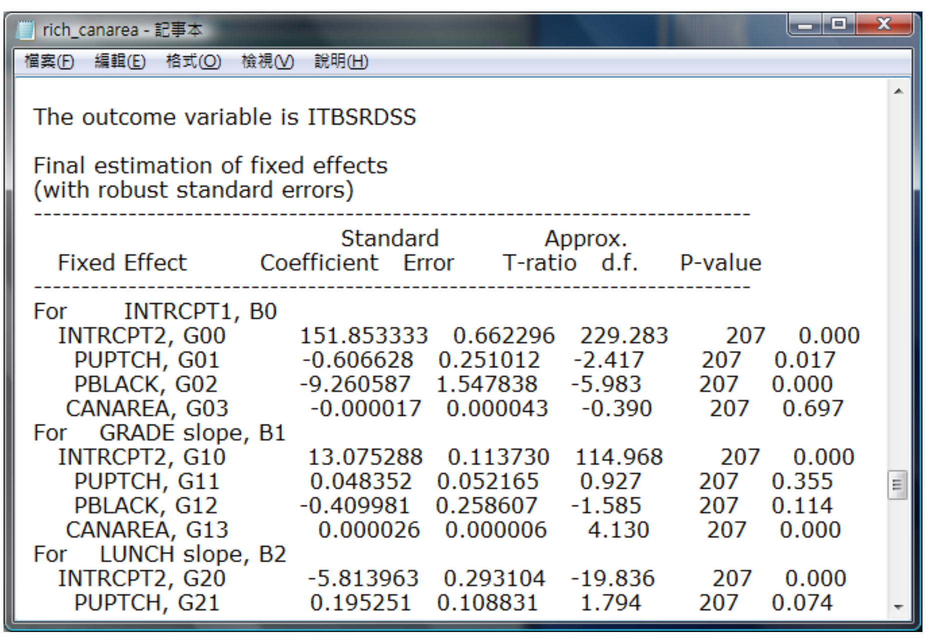


7. After specifying level-1 model, we can specify level-2 model.



8. After specifying the models, click "Run Analysis" from the tool bar above. It will take HLM some time to iterate the MLE solution. When it is done, click "File" from the tool bar above and click "View Output."

9. You will find the report as we see before.



rich_canarea - 記事本

檔案(F) 編輯(E) 格式(O) 檢視(V) 說明(H)

The outcome variable is ITBSRDSS

Final estimation of fixed effects
(with robust standard errors)

Fixed Effect	Coefficient	Standard Error	T-ratio	Approx. d.f.	P-value
For INTRCPT1, B0					
INTRCPT2, G00	151.853333	0.662296	229.283	207	0.000
PUPTCH, G01	-0.606628	0.251012	-2.417	207	0.017
PBLACK, G02	-9.260587	1.547838	-5.983	207	0.000
CANAREA, G03	-0.000017	0.000043	-0.390	207	0.697
For GRADE slope, B1					
INTRCPT2, G10	13.075288	0.113730	114.968	207	0.000
PUPTCH, G11	0.048352	0.052165	0.927	207	0.355
PBLACK, G12	-0.409981	0.258607	-1.585	207	0.114
CANAREA, G13	0.000026	0.000006	4.130	207	0.000
For LUNCH slope, B2					
INTRCPT2, G20	-5.813963	0.293104	-19.836	207	0.000
PUPTCH, G21	0.195251	0.108831	1.794	207	0.074