Using Multiple Regression to Ascertain Group Differences in the Relationship of Predictors to a Criterion:

Ethnic Group Differences in the Relationship between Course-taking and Achievement in Mathematics

A DISSERTATION
SUBMITTED TO THE FACULTY OF THE
UNIVERSITY OF MINNESOTA

BY

Kyungin Park

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY

Ernest C. Davenport Jr.

October 2018

Acknowledgment

First, I would like to express my sincere gratitude to my academic advisor, Prof. Ernest Davenport Jr., for his support, guidance, and for believing in me even when I had little belief in myself. I am also extremely grateful to the rest of my thesis committee members, Prof. Mark Davison, Prof. Robert delMas, and Dr. Suh-Ruu Ou, for their support and encouragement.

My appreciation also goes to all my friends, for each and everyone one of them, for their care, love, emotional support, and presence.

> Family-like friends in MN (the Parks, the Lees, the Yongs, and the Yoos) and the St. Andrew Kim Catholic church community were my family in MN and always made me feel so loved.
>
> I was able to endure this journey owing to my bestie Sangyoo, heartful Somie, reliable Sooyeon, smart Yon Ji, hot Su Jung, cheerful Ehwa, faithful Hyun, helping Sangok, hopeful Sehyun, admirable Kyle, brilliant Jean-Baptiste, sincere Liz, adorable Bowen, trustful Kory, delightful Jonathan, warm Astrid, genius Christina, and all other friends for caring, supporting, and spending precious time with me.
>
> I sincerely thank Jane for encouraging me like a mom, teaching me how to be kind to myself as well as to take good care of myself.
>
> Special thanks to Young Sun Kwon esq., checking in every morning in spite of her busy schedule, encouraging me, and always sharing little moments in life far away in D.C. like sisters.

I am very much indebted to my loving family and my in-laws, I deeply appreciate their support and patience.

> I am so much grateful to my parents Kwang Myung and Jeong Soo, my parents' in-laws Yong Kil and Young Sun, my sister Kyung Jin, my brothers Kyoung

Hwan and Kyoungjune, sister in-law Jio, and their spouses and kids for their endless love and support.

Hyeongryeol, my husband, I cannot thank you enough for your love and support both inside and out, and always wanting the best for me. Your support and encouragement convinced me that I have the strength to finish this journey and can be a cafe owner later in life.

I started my PhD program in Korea about ten years ago, so it has been quite a long journey to put three letters after my name, PhD.

Thank you, God. I'm PHinally Done!

Abstract

This study introduces an approach that utilizes ten regression models to assist in the study of the relationship between predictors and a criterion as it pertains to group differences. The models range in complexity from a simple multiple regression model that assumes one size fits all, to a full moderated multiple regression model, where each group has its own set of predictors. The different models arise via varying constraints that are theoretically meaningful and testable relative to equity of the groups. Data for mathematics achievement and course-taking for different ethnic groups obtained from the National Center for Education Statistics' (NCES) High School Longitudinal Study of 2009 (HSLS: 09) are used to illustrate and explain the different models. This study is especially timely given the attention to achievement gaps and their causes and correlates.

Table of contents

List of Tables……………………………………………………………………………vii

List of Figures…………………………………………………………………………… x

I. Introduction 1

II. Review of Literature 8

A. Methodological Review 8

i) Moderated regression model approach 8

ii) Ten models for assessing group differences in relationships 16

iii) *Level* and *Pattern* 33

B. Substantive Review: course-taking and achievement in mathematics 37

C. Synopsis 39

III. Methodology 42

A. Data 42

B. Measures 43

i) Independent variables 43

ii) Moderator variable 45

iii) Dependent variable 46

iv) Covariates 47

C. Methods 49

i) Ten multiple regression models 49

ii) Model comparison 50

iii) Statistical program 53

iv) Design effect 55

IV. Results 57

A. Descriptive statistics and correlation 57

B. Ethnic differences in the relationship between thirteen sequences of mathematics course-taking and mathematics achievement 64

i) Base model 64

ii) ANCOVA 66

iii) PMM 68
iv) MMM 71
v) LMM 74
vi) FMM 76
vii) Model Comparisons 81
C. Ethnic differences in the relationship between *Level* and *Pattern* of mathematics course-taking and mathematics achievement 85
i) Base model 85
ii) ANCOVA 86
iii) PMM 87
iv) MMM 89
v) LMM 91
vi) FMM 93
vii) Model Comparisons 95
D. Possible covariates: prior achievement and socio-economic status (SES) 99
i) Prior achievement 99
ii) Socio-economic status (SES) 100
E. Ethnic differences in the relationship between prior achievement, thirteen sequences of mathematics course-taking and mathematics achievement 103
i) Base model 103
ii) ANCOVA 107
iii) PMM 108
iv) MMM 111
v) LMM 114
vi) FMM 117
vii) Model Comparison 123
F. Ethnic differences in the relationship between prior achievement, *Level* and *Pattern* of mathematics course-taking and mathematics achievement 127
i) Base model 127
ii) ANCOVA 128
iii) PMM 129
iv) MMM 131
v) LMM 133

vi) FMM 135
vii) Model Comparison 137
G. Testing the assumption for the ANCOVA model 142
V. Discussion 144
VI. Conclusion 152
References 157

List of Tables

Table 3.1
Frequency table and average post achievement scores by ethnic group 49
Table 4.A.1
Descriptive Statistics with transcript weights for overall sample (n=15,750[*]) 58
Table 4.A.2
Descriptive Statistics with transcript weights by ethnic group 59
Table 4.A.3
Correlations among prior and post mathematics achievement, Level, and Pattern for overall and by ethnic group with transcript weights 61
Table 4.A.4
Number of Carnegie Units in Each Sequence and Other Scores for Five Students 62
Table 4.B.1
Results for the base model predicting mathematics achievement 65
Table 4.B.2
Results for the ANCOVA model predicting mathematics achievement 67
Table 4.B.3
Results for PMME predicting mathematics achievement 69
Table 4.B.4
Results for PMMU predicting mathematics achievement 70
Table 4.B.5
Results for MMME predicting mathematics achievement 71
Table 4.B.6
Results for MMMU predicting mathematics achievement 73
Table 4.B.7
Results for LMME predicting mathematics achievement 74
Table 4.B.8
Results for LMMU predicting mathematics achievement 75
Table 4.B.9
Results for FMME predicting mathematics achievement 77
Table 4.B.10
Results for FMMU predicting mathematics achievement 79
Table 4.B.11
Model fit values for each model 83
Table 4.B.12
Results for F-tests of Model comparison 84
Table 4.C.1
Results for the base model predicting mathematics achievement 86
Table 4.C.2
Results for the ANCOVA model predicting mathematics achievement 87
Table 4.C.3
Results for PMME predicting mathematics achievement 88
Table 4.C.4
Results for PMMU predicting mathematics achievement 89

Table 4.C.5
Results for MMME predicting mathematics achievement 90
Table 4.C.6
Results for MMMU predicting mathematics achievement 91
Table 4.C.7
Results for LMME predicting mathematics achievement 92
Table 4.C.8
Results for LMMU predicting mathematics achievement 93
Table 4.C.9
Results for FMME predicting mathematics achievement 94
Table 4.C.10
Results for FMMU predicting mathematics achievement 95
Table 4.C.11
Model fit values for each model 97
Table 4.C.12
Results for F-tests of Model comparison 97
Table 4.D.1
Regression analyses predicting mathematics achievement using Level and Pattern (Model 1), and adding Prior Achievement (Model 2) 99
Table 4.D.2
Regression analyses predicting mathematics achievement using prior achievement, Level and Pattern (Model 1), and adding SES (Model 2a) or SES with locale (Model 2b) 102
Table 4.E.1
Results for the base model predicting mathematics achievement 104
Table 4.E.2
Results for the ANCOVA model predicting mathematics achievement 108
Table 4.E.3
Results for PMME predicting mathematics achievement 109
Table 4.E.4
Results for PMMU predicting mathematics achievement 110
Table 4.E.5
Results for MMME predicting mathematics achievement 112
Table 4.E.6
Results for MMMU predicting mathematics achievement 114
Table 4.E.7
Results for LMME predicting mathematics achievement 115
Table 4.E.8
Results for LMMU predicting mathematics achievement 116
Table 4.E.9
Results for FMME predicting mathematics achievement 118
Table 4.E.10
Results for FMMU predicting mathematics achievement 121
Table 4.E.11
Model fit values for each model 125
Table 4.E.12
Results for F-tests of Model comparison 125

Table 4.F.1
Results for the base model predicting mathematics achievement 127
Table 4.F.2
Results for the ANCOVA model predicting mathematics achievement 128
Table 4.F.3
Results for PMME predicting mathematics achievement 130
Table 4.F.4
Results for PMMU predicting mathematics achievement 130
Table 4.F.5
Results for MMME predicting mathematics achievement 131
Table 4.F.6
Results for MMMU predicting mathematics achievement 132
Table 4.F.7
Results for LMME predicting mathematics achievement 133
Table 4.F.8
Results for LMMU predicting mathematics achievement 134
Table 4.F.9
Results for FMME predicting mathematics achievement 135
Table 4.F.10
Results for FMMU predicting mathematics achievement 136
Table 4.F.11
Model fit values for each model 140
Table 4.F.12
Results for F-tests of Model comparison 140
Table 4.G
Results for regression predicting mathematics achievement using ethnic group 142

List of Figures

Figure 1. Moderator model (left) and mediator model (right) adapted from “The Moderator-Mediator Variable Distinction in Social Psychological Research: Conceptual, Strategic, and Statistical Considerations.” by R. M. Baron and D. A. Kenny, 1986, *Journal of Personality and Social Psychology, 6*, p. 1174 and p. 1176. 11

Figure 2. Unrestricted interrupted model (left) and Spline (Restricted interrupted) model (right) of the U.S. population before and after World War II. Adapted from “Quantitative applications in the social sciences: Spline regression models” by L. C. Marsh and D. R. Cormier, 2002. 14

Figure 3. All ten models reviewed in this thesis (the numbers in front of model names are matched with the numbers for each equation below) 17

Figure 4. Four families of analyses (B, C, E and F sections in Chapter IV) with two types of predictors and with or without the presence of a covariate (prior achievement) 48

Figure 5. Unstandardized regression coefficients to predict mathematics achievement with thirteen course-taking categories for overall students 66

Figure 6. Unstandardized regression coefficients for the base model for predicting mathematics achievement with *Level* and *Pattern* (left) vs. Prior achievement, *Level*, and *Pattern* (right) for overall and each ethnic group 146

I. Introduction

The achievement gap is defined by a disparity in measures of educational performance between subgroups of students (e.g. ethnicity, gender, socioeconomic status, etc.). Research and efforts to describe and narrow achievement gaps have been around for a long time. Coleman (1961) used disparities in standardized test scores as indices of unequal opportunity and concluded that a student's family background was the biggest determinant of how well a student would learn. The Coleman Report succeeded in starting discussions on educational policy; we are still questioning today how interventions might help reduce gaps and improve student outcomes. It is timely and critical that scholars develop statistical methods to assist in this research effort. Understanding the causes and correlates of achievement gaps will help develop methods and interventions aimed at decreasing achievement gaps. In response to this need, the current research was designed to further understand causes and correlates of achievement gaps. Specifically, mathematical models suggested in this thesis, can empirically test theoretical differences in relationships between groups.

Many studies have focused on student-level factors, including their course-taking (Bozick & Ingels, 2008; Davenport et al., 2013; Davenport, Park, Ayodele, Davison, & Kang, 2015; Riegle-Crumb & Grodsky, 2010), attitudes and aspirations (Hastedt, 2016; Son, Watanabe, & Lo, 2017), academic motivation (Son et al., 2017), self-efficacy (Yeh, 2017), socio-economic background, and gender and ethnicity (Bohrnstedt, Kitmitto, Ogut, Sherman, & Chan, 2015; Cameron, Grimm, Steele, Castro-Schilo, & Grissmer, 2015; Davis-Kean & Jager, 2014; Hoffer, Rasinski, & Moore, 1995; Peterson, Rubie-

Davies, Osborne, & Sibley, 2016; Yeh, 2017; Young, 1994). Specifically, Black and Hispanic students are less likely to enroll in advanced mathematics courses before graduation compared to White and Asian students (Bozick & Ingels, 2008; Dalton, Ingels, Downing, & Bozick, 2007; Davenport et al., 1998). This might be related to the fact that Black and Hispanic students tend to have lower mean achievement scores prior to high school than White and Asian students.

In addition to categorizations within students, cultural and structural factors, such as home, community, and in-school factors, are related to academic performance and also contribute to achievement gaps, especially in mathematics (Cameron et al., 2015; Davis-Kean & Jager, 2014; Hoffer et al., 1995; Peterson et al., 2016; Vanneman, Hamilton, Baldwin Anderson, & Rahman, 2009; Yeh, 2017). School-level factors and teacher or classroom-level factors, for example, school resources and climate (Bohrnstedt et al., 2015; Bozick & Ingels, 2008; Hastedt, 2016; Reeves, 2012) and teachers' expectations and attitudes toward students (Peterson et al., 2016) also contribute to the achievement gap. These achievement gaps were found to be related to success in higher education and future life factors, including life satisfaction and future income (Reardon, 2013).

Previous research has employed multiple regression using dummy variables for groups, but these models do not fully account for the different relationships between predictors and a criterion between groups. For example, one group might have experienced a particular intervention that changes the relationship between X and Y (e.g.

implementing a special program for low-income students changed the relationship between their class participation and achievement scores). In this case, one may specify conditions (Z) that strengthen or weaken the relationship between the X(s) and Y. For example, the effect of the same policy may be associated with the criterion differentially between groups due to differences in experiences or features those groups possess. If the relationships are the same between groups, one does not need to use a moderated regression model since it lacks efficiency. In fact, many studies have focused on relationships between predictors (e.g., school engagement, educational policy, course-taking, academic motivation, level of mother's education) and the criterion (e.g., mathematics achievement, school satisfaction) to address narrowing achievement gaps between groups with different characteristics (e.g., ethnicity, gender, and English language learners). For instance, when the relationship between mathematics course-taking and mathematics achievement is stronger for group A than for group B, then group A might have a larger increment in the expected Y (mathematics achievement) than group B.

There have been studies to model such situations (Cook & Weisberg, 2004; Davison, Davenport, and Kohli, 2017; Saunders, 1956), but few have proposed and reviewed specific models and examined systematic differences among them. These models range from a simple multiple regression model, which assumes one set of predictors applies to all groups, to a full moderated regression where each focal group has its own set of predictors. When differences between groups exist in relationships between predictors and a criterion, one may assume that there are differences in slopes and intercepts between a reference group and focal groups and propose a full moderated regression model. The moderated regression model predicts differences in parameters between a reference and a focal group and uses adjustments for the focal group compared to the reference group. If one size fits all, the base model is sufficient, and adjustments for focal groups are not needed. In this case, the base model with one intercept and one set of slopes for all groups is sufficient.

The purpose of this thesis is to introduce an approach that uses regression models to study group variation when there are differences between groups in the relationships between predictors and a criterion. These models begin with a simple one-size-fits-all regression model and builds to the full moderated regression model. The intermediate models are subsets of the full moderated model in which the constraints are both testable and theoretically meaningful relative to group differences. The existing literature does not review all ten models discussed here and this study addresses this limitation of research in order to analyze the important topic of achievement gaps between ethnic groups.

In this thesis, four new models will be introduced, the Mean Moderated Model (MMM) with or without varying intercepts and the Linearly Moderated Model (LMM) with or without varying intercepts, which add to the literature of moderated regression

(Saunders, 1956). A partial one-dimensional model (Cook & Weisberg, 2004) will also be explored. MMM and LMM also apply multiple regression methods to study group variation when groups seem to have different relationships between predictors and a criterion.

After systemically reviewing ten models, including two families of new intervening models (MMM with or without equal intercept and LMM with or without equal intercept), a case study using real data will be presented. The best model when using these data is also suggested to show how its mathematical constraints relate to theoretical aspects of mathematics course-taking and mathematics achievement for different ethnic groups. The High School Longitudinal Study of 2009 (HSLS:09) data are analyzed to demonstrate the way these models can be used to understand the differences in the relationship between course-taking and achievement between groups. In addition, a computer program to implement all proposed models including PMM, MMM, and LMM is also provided. The computer program will allow one to account for weights and design effects.

When analyzing real data to explore differences in relationship between course-taking and achievement among ethnic group, *Level* and *Pattern* of course-taking were also used as predictors to examine whether the amount or the content of course-taking is more related to mathematics achievement. According to Davison and Davenport (2002), all the predictive information contained in a group of predictors ($v \geq 2$) can be reparametrized into two variables, *Level* of the predictors and *Pattern* of the predictors. For example, credits completed in mathematics courses, can be represented in two components: *Level* (the total or mean number of completed credits in mathematics courses) and *Pattern* (the covariance-like statistic, having a distribution of course-taking related to high mathematics achievement) of the courses (Davison & Davenport, 2002). These two predictors carry all the predictive information about the amount of course-taking (*Level*) and the contents of course-taking (*Pattern*), and they explain both quantitative and qualitative features of mathematics course-taking.

This thesis will attempt to answer the following research questions:

1. Given a general linear model approaches, specifically subsets of full moderated multiple regression:
    a. How can ten subsets of a moderated regression model based on slope and intercept constraints be used to investigate group differences in regression weights?
    b. How can those constraints be explained in a theoretically and statistically meaningful and testable way?
    c. How can these models be compared and constraints be tested relative to group differences?
2. How do the relationships between mathematics course-taking and mathematics achievement differ between ethnic groups?

a. Among the ten models, which model is optimal when using real data?

b. How can these models be interpreted?

c. How do credits earned in thirteen sequences distinctly relate to mathematics achievement between ethnic groups?

d. Do *Level* and/or *Pattern* of mathematics courses taken differentially relate to mathematics achievement for four ethnic groups?

II. Review of Literature

In this chapter, both the methodological part and the substantive part of this study will be reviewed. In the methodological review, the general regression approach and moderated multiple regression are introduced first. Then, an integrative review of ten models for assessing group differences in relationships is suggested.

A. Methodological Review

i) Moderated regression model approach

a) Introduction to linear modeling

Linear models are a mainstay of statistical methods in the social sciences with regression models being among the most popular. Investigators have tried to find a parsimonious way to model behaviors or phenomena, and many research questions have as their premise linear relationships among variables. Among the simplest models is the linear model, which implies a linear relationship exists between a predictor (independent variable, X) or predictors (Xs) and a criterion (dependent variable, Y). In other words, the additive effect of one unit increase in an independent variable(s) on a dependent variable is constant. Additionally, the additive effect is the same regardless of subpopulation since the linear model assumes that there is only one population (Friedrich, 1982).

b) Moderated regression

Moderated regression (Saunders, 1956) is based on the concept of interaction regression modeling (Aiken & West, 1991). Kam and Franzese (2007) describe interactive models as “second generation models” due to their increased complexity and ability to parse whether results obtained from simple linear models are general versus conditional on some third variable (the moderator). Social scientists have often evaluated such hypotheses using linear interactive or multiplicative terms (Kam & Franzese, 2007). They attempt to specify conditions (Z) that strengthen or weaken the relationship between X and Y. For example, the effects of some set of individual characteristics (e.g., ideology or academic motivation) on a criterion related to another set of individual characteristics (e.g., gender or ethnicity), and these can and should be analyzed with interactive terms. Research questions that ask how the impact of some experimental treatment or environmental context (e.g. participation in an ACT/SAT preparation program or having a high score on a SES index) relates to the level of some individual characteristic (e.g. academic motivation or supportive home environment) likewise imply interactive hypotheses. For instance, when the same regression equation is used for both men and women, it is found that the predicted grade for women is underestimated relative to that for men who earn equal actual grades (Stricker, Rock, & Burton, 1993). In short, women’s predicted grade is underestimated because the model used relies on the same regression equation whereas gender actually works as a moderator.

Saunders (1956) introduced moderated multiple regression (MMR), which determines the relationship between Y and Xs, moderated by a third variable Z. In his original paper, the parameter Z did not indicate membership in a group, but a score on some continuous variable (Saunders, 1956). Gaylord and Carroll (as cited in Zedeck, 1948) called the moderator variable a "population control variable," which identifies sub-populations in which the application of a multiple regression equation for the entire population is inappropriate. Grooms and Endler (1960) used a "modifier variable," which dichotomized or trichotomized Z to differentiate subgroups, and they point out that it is different from the moderator Saunders (1956) proposed. Banas (1964) claims that all the definitions above imply some form of interaction between variables and subgroups. While the original moderator model considered the moderator to be a quantitative, continuous variable (Saunders, 1956), the procedure operates the same whether the moderator is continuous or categorical (Banas, 1964). This study provides various moderated models to specify group membership (g=0, 1, 2, ..., G-1) when the moderator (e.g. ethnicity) is categorical. Specifically, this study focuses on interactions between a grouping factor and continuous predictors, but not interactions between the continuous predictors (Cook & Weisberg, 2004).

c) Moderators vs. Mediators

Although not necessarily mutually exclusive, it is important to distinguish a moderating from a mediating effect since they differ conceptually, strategically, and statistically (Aguinis, 2004; Baron & Kenny, 1986). A mediator variable, which is also called an "intervening" or a "process" variable, accounts for the causal relation between *X* and *Y*. It gives information on why or how such an effect occurs, whereas a moderator variable specifies when or under what conditions certain effects hold (Aguinis, 2004).

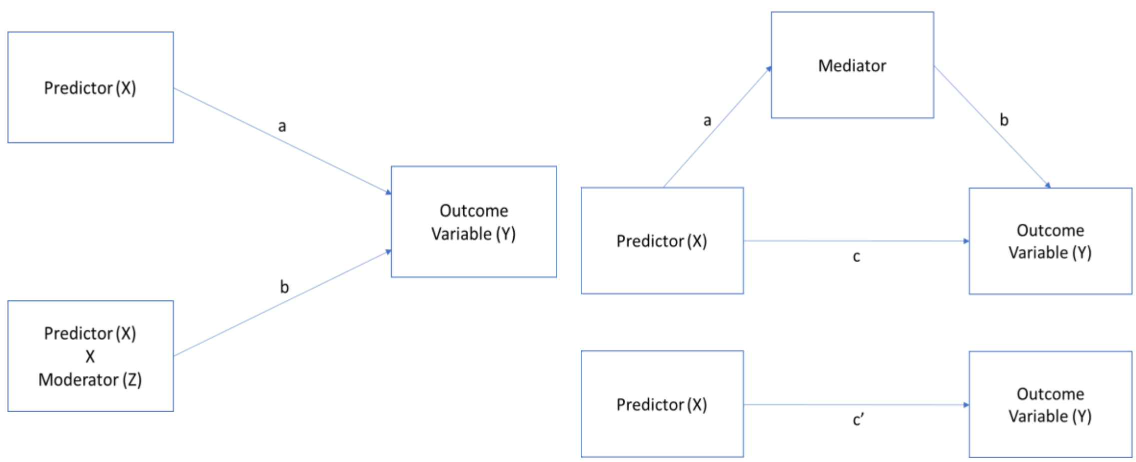
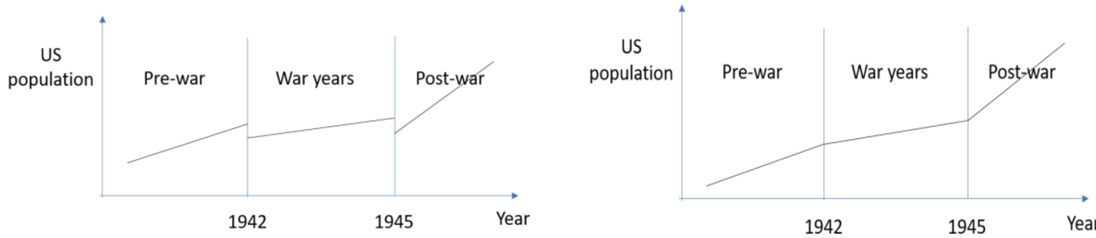


Figure 1. Moderator model (left) and mediator model (right) adapted from "The Moderator-Mediator Variable Distinction in Social Psychological Research: Conceptual, Strategic, and Statistical Considerations." by R. M. Baron and D. A. Kenny, 1986, *Journal of Personality and Social Psychology, 6*, p. 1174 and p. 1176.

The mediator model is presented in Figure 1 (right). In this conception, one should estimate the three following regression equations to test the mediation: regressing the mediator on the predictor (path **a**), regressing the outcome variable on the predictor (path **c’**), and regressing the outcome variable on both the predictor and the mediator (paths **b** and c). To establish mediation, three conditions must be met. The predictor significantly accounts for variation in the mediator (path a), the mediator significantly accounts for variation in the outcome variable (path b), and the effect of the predictor on the outcome variable when paths a and b are controlled in the predicted direction (path c) must be less than the effect of the predictor on the outcome variable (path c’). The Stimulus-Organism-Response (S-O-R) model is one example of a mediator model: it recognizes that an active organism intervenes between stimulus and response (Woodworth, as cited in Baron and Kenny, 1928). Here, the stimulus is the predictor, the response is the outcome variable, and the active organism (various transformation processes) is the mediator. The main idea of this model is that the effects of stimuli on behavior are mediated by various transformation processes internal to the organism. If the predictor has no effect on the outcome variable (path c=0) when the mediator is in the model (controlled for), perfect mediation holds (Baron & Kenny, 1986). In other words, perfect mediation holds when stimulus is related to response, but stimulus does not account for response when both stimulus (a predictor) and organism (the mediator) are included in the model. For instance, behavioral inhibition (stimulus) effects on stress reactivity (response) only occur via secure attachment (the mediator, the active organism).

d) Spline models

When there are group differences, one may also use spline regression models instead of moderated regression models. Marsh and Cormier (2002) describe intervention/interrupted models, spline models (including piecewise linear models), and dummy variable models as being very similar forms of regression analysis. For instance, spline regression models have two or more segments of lines without any breaks or jumps. These lines change their direction at points called spline knot(s), where they join. The spline regression model is a *restricted* interrupted model, where the functions are forced to be continuous at the knot(s) with possible changes in slopes. An *unrestricted* interrupted model, however, may have different intercepts and slopes among lines. For example, the regression line for the dependent variable of interest may suddenly change its slope (a kink in the line) without causing a break or a jump in the line itself. For instance, one may predict the US population over three time periods (pre-World War II, during World War II, and after World War II), which is shown in Figure 2. Using an *unrestricted* interrupted model, separate regression lines with different slopes and intercepts for the US population would be suggested (left). Using a spline regression model (*restricted* interrupted model, right), the function is forced to be continuous at two spline knot locations, the years 1942 and 1945, whereas its slope may be discontinuous (Marsh & Cormier, 2002).



Unrestricted interrupted model Spline (Restricted interrupted) model

Figure 2. Unrestricted interrupted model (left) and Spline (Restricted interrupted) model (right) of the U.S. population before and after World War II. Adapted from "Quantitative applications in the social sciences: Spline regression models" by L. C. Marsh and D. R. Cormier, 2002.

As suggested by the name, the dummy variable regression model contains a dummy variable (also known as an indicator, categorical, binary, or qualitative variable), which can take the value 1 or 0 to indicate the presence or absence of some categorical effect that may be expected to shift the outcome. For example, a gender (female) scored 0 for male and 1 for female, X is academic motivation, and Y is mathematics achievement. $Y = \beta_0 + \beta_1 * X + \beta_2 * (female) + \beta_3 * (female) * X + e$, produces two equations, one for each group (male and female) where β_2 is an adjustment for the intercept for the female group and β_3 is an adjustment for the regression weight for the female group (these adjustments do not appear in the regression equation for males). This dummy variable regression model is one example of a moderator model. For instance, when academic motivation is a predictor, gender is a moderator, and the achievement score is an outcome variable, the dummy variable regression model with a dummy variable (here, *female*) works like a moderator model. Additionally, interrupted regression uses a dummy variable for 'intervention' to show whether or not the intervention was implemented. The change can be either gradual or abrupt in intervention/interrupted regression methods, whereas splines are defined as functions that are continuous and focus on smoother transitions at spline knots (Marsh & Cormier, 2002).

There are three important features in applying spline methods: the number of distinct spline segments into which the predictor variable falls (e.g. year), the degree of the polynomial used to represent each segment, and the location of spline knots (Marsh & Cormier, 2002). The simplest case is where all three features are assumed to be known. The most complex case happens when the location of the spline knots is unknown, and the number of segments is also unknown. Spline modeling is one way of fitting data, and the main difference in comparison with the moderated models is that spline modeling is not a priori. Therefore, when one needs to build a model with a priori theory, moderated multiple regression models are useful because there may be differences in the relationships between predictors and the criterion between groups. In addition, moderated multiple regression modeling can represent all groups using different parameters in one equation. In other words, moderated multiple regression models consist of one equation with parameters for each focal group adjustment, and so are simpler than having to specify many separate equations.

ii) Ten models for assessing group differences in relationships

There are plenty of possible linear models between the simplest (1) base model (simple multiple regression model) and the (2) full moderated multiple regression model (FMM model; Saunders, 1956) with respect to parameters, but this study specifically focused on intervening models that have regular patterns in regression coefficients that have theoretical implications relative to the predictors, the criterion, and the moderator. The first intervening model is the Partial One-Dimensional Model (POD model; Cook & Weisberg, 2004), also called the Proportional Moderated Model (PMM); the second is the Mean Moderated Model (MMM); and the third is the Linearly Moderated Model (LMM; a combination of the PMM and the MMM models). These three intervening models are more parsimonious than the full moderated model, but more complex than the base model. Each of these three models can allow intercepts to vary or force intercepts to be constant, and so they yield six models. Adding to the above six possible models, one can force intercepts to be constant in the FMM, which can be called FMME (FMM with an Equal intercept) or the restricted full moderated model. Plus, the ANCOVA model appears when one allows the intercepts to vary across groups from the base model. Therefore, a total of ten possible models to explain group differences in relationships will be addressed (See Figure 3).

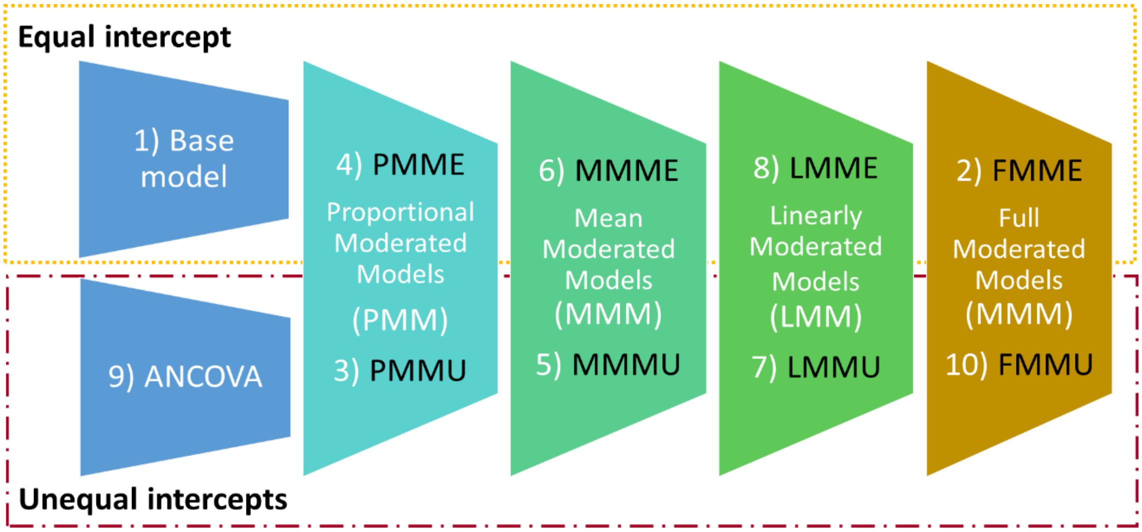


Figure 3. All ten models reviewed in this thesis (the numbers in front of model names are matched with the numbers for each equation below)

a) Base model (one size fits all)

The base model is just the multiple regression model assuming that there are no group differences in relationship to a criterion. The base multiple regression model aggregates all values as if the participants are from one population without considering subgroups (Equation 1).

$$Y = \sum_v \beta_v X_v + \beta_0 + e \tag{1}$$

$$DF_1 = V, DF_2 = N - (V + 1)$$

Here, β_v is the least squares estimate of the population regression coefficients for v continuous variables (v=1, 2, …, V) common to all groups, β_0 is the least squares estimate of the population intercept, and e is an error term. In short, Equation 1 is just a simple multiple regression model. There are the same slopes and the same intercept for all groups (one size fits all) under the assumption that there are no group differences. If this model is the one chosen from model comparison, there is no statistical difference in the relationship between the predictors and the criterion for subgroups, and the moderator has no effect.

b) Full moderated regression model (FMM)

Equation 2 is the full moderated multiple regression model (FMM), which is the most complex model among all suggested models in this study.

$$Y = \sum_v \beta_{v0} X_v + \sum_{g=1}^{g=G} \sum_v \beta_{vg} z_g X_v + \sum_{g=1}^{g=G} \beta_{2g} z_g + \beta_0 + e \qquad (2)$$

$$DF_1 = (G+1)V + G,\ DF_2 = N - (G+1)(V+1)$$

Here, β_{v0} is the regression coefficient for variable v in the reference group (g = 0), β_{vg} is the difference between the regression coefficient for variable v in focal group g and the regression coefficients for the reference group (β_{v0}). The term z_g is an indicator variable which is 1 for focal group g and 0 for all other groups, and β_{2g} is the difference between the intercept in focal group g and the intercept for the reference group. β_0 is the intercept in the reference group, and $(G+1)$ is the number of total groups including the reference group ($g = 0$).

If F-tests are significant when comparing the FMM with all other models, one can conclude that there are differences in the relationships between the predictors and the criterion between groups, and these differences are not depended on additive and/or proportional constraints on regression coefficients. This full model has the largest number of parameters to be estimated among all ten models suggested in this study.

c) Proportional Moderated Model (PMM)

Cook and Weisberg (2004) developed a restricted model that is more general than the base model, but less general than the full moderated model. They called it the Partial One-Dimensional (POD) Model, but Proportional Moderated Model (PMM) explains this model better. This model allows for the difference in vectors of regression weights to vary across groups but only by a constant of proportionality: $\beta_{vg} = k_g \beta_{\boldsymbol{v0}}$ where $\boldsymbol{\beta}_{vg}$ is the difference between vectors of the regression weights between a focal group g and the reference group, k_g is the constant of proportionality for each focal group g, and $\boldsymbol{\beta_{v0}}$ is a vector of the regression weights in the reference group. Note that the constant of proportionality can differ for each group.

$$Y = \sum_v \beta_{v0} X_v + \sum_{g=1}^{g=G} \sum_v k_g \beta_{v0} z_g X_v + \sum_{g=1}^{g=G} \beta_{2g} z_g + \beta_0 + e \qquad (3)$$

$$DF_1 = V + 2G, DF_2 = N - (V + 2G + 1)$$

Cook and Weisberg (2004) proposed that this model is more interpretable when the usual analysis of covariance model is unacceptable, and the full moderated model is more complex than necessary to account for the data. The Proportional Moderated Model with an Equal intercept (PMME) allows the vector of regression weights to vary across groups by a constant of proportionality, while constraining the intercept to be constant for all groups (See Equation 4).

$$Y = \sum_v \beta_{v0} X_v + \sum_{g=1}^{g=G} \sum_v k_g \beta_{v0} z_g X_v + \beta_0 + e \quad (4)$$

$$DF_1 = V + G, DF_2 = N - (V + G + 1)$$

When comparing the base model and PMME, one can test the null hypothesis, H_0: $k_g = 0$ for all $g \geq 1$, while the alternative hypothesis would be H_a: $\beta_{vg} = k_g \times \beta_{v0}$ for all (v, $g \geq 1$) and $k_g \neq 0$ for at least one group $g \geq 1$. Here, *v* denotes the regression weights for predictors and this *F*-test allows one to test the proportionality of different regression weights across groups. If this *F*-test is significant, one size does not fit all groups and it is useful having differences in slopes (the relationship between the predictors and the criterion) which are proportional ($k_g \beta_{v0}$). If the *F*-test is not significant, however, the simpler model is more parsimonious, which means relationships between predictors and a criterion do not differ across groups (one size fits all).

When comparing PMM and FMM, one can test the null hypothesis, H_0: $\beta_{vg} = k_g \times \beta_{v0}$ for all (v, g ≥1) while the alternative hypothesis would be H_a: $\beta_v \neq k_g \beta_{v0}$ for some (v, g ≥1). This test allows for testing whether PMM is too restrictive and we need the full moderated model because the regression weights for some of the groups are not multiplicatively different. If the *F*-test is not significant when comparing PMM with FMM, FMM has some unnecessary parameters being estimated, and a proportional value can represent the differences in all other regression weights between groups. If the *F*-test is significant, PMM is too restricted and FMM is needed because the regression weights for some of the groups does not simply differ by a multiplicative constant.

When all k_gs are equal to 0, this PMM becomes the ANCOVA model, which has the same regression weights vector (β_{v0}) for all groups. When k_g is -1, the regression weight for that specific focal group g equals to 0. When k_g equals to 1, the difference in slopes between focal group g and the reference group is β_{v0}, and the slope for the focal group g is $2\beta_{v0}$. Moreover, if k_g is greater than zero, we know that the focal group in question has a larger increment in the expected Y (criterion) than the reference group. When k_g is greater than -1 and less than 0, the focal group in question has a smaller increment in the expected *Y* than the reference group. PMM has some useful properties. The rank order of the absolute value of each population's regression weights (v=1, 2, …, V) remains the same in all subgroups. In addition, the difference vector (for each focal group) of all regression weights changes sign in parallel, since the differences in all regression weights are the product of these weights multiplied by k_g for each focal group. In other words, the order of the impact or strength of the relationship for predictors (the order of the absolute value of regression weights) always stays the same, but the direction of the relationship for focal group g could be changed when k_g is less than -1.

We will use the following example to explain all the models: the predictors are credits in five courses (Algebra 2, Geometry, Statistics, Pre-Calculus, and Calculus) students have completed in mathematics, mathematics achievement score is the criterion, ethnic group is the moderator (z_g), and the White group is the reference group. When k_{Asian} is greater than 0, the Asian group has a larger increment in the expected Y (criterion) than the White group. In other words, all predictors are associated with a larger change in math achievement for the Asian group of students than the White group of students when $k_{Asian}>0$. This k_g is a different constant for each focal group g, and PMM gives different parameters, which differ by proportionality.

d) Mean Moderated Model (MMM)

Davison, Davenport, and Kohli (2017) proposed another model, the Mean Moderated Model (MMM) intermediate between (1) and (2), a model in which the regression weights for focal group g and the reference group differ by an additive constant β_g^*. This model allows regression weights for any two groups to differ, but the difference in regression weights is constrained such that the difference is an additive constant (β_g^*) for each group, g across variables.

Like PMM (Cook and Weisberg, 2004), the Mean Moderated Model is also a special case of the full moderated model. This model for any focal group (g ≥1) has the following equation:

$$Y = \sum_v(\beta_{v0} + \beta_g^*)X_v + \beta_{0g} + e = \sum_v \beta_{v0}X_v + \sum_v \beta_g^* X_v + \beta_{0g} + e$$
$$= \sum_v \beta_{v0}X_v + \beta_g^* \sum_v X_v + \beta_{0g} + e \qquad (5\text{-}1)$$

For each focal group with $g \geq 1$, all regression weights for the predictors are within an additive constant β_g^* of the corresponding weights for the reference group. If coefficients for the two groups differ only by an additive constant, the regression weight vectors for the two groups will differ only in their means; hence, it is called the Mean Moderated Model (hereafter MMM). $\sum_v \beta_{v0}X_v$ is a function of the regression coefficients and predictors in the reference group, and $\beta_g^* \sum_v X_v$ is a single predictor variable, in which $\sum_v X_v$ is multiplied by a regression weight β_g^*. Then, $\sum_v X_v$ is just the sum of the predictors for one observation, and it is referred to as the "total score predictor variable" (Davison et al., 2017). This linear equation (5-1) has *V+1* predictors, which are *V* predictors (X_v) and the total score ($\sum_v X_v$).

This MMM with G focal groups can be generalized using the notations as for FMM and PMM:

$$Y = \sum_v \beta_{v0} X_v + \sum_{g=1}^{g=G} \beta_g^* z_g \left(\sum_v X_v\right) + \sum_{g=1}^{g=G} \beta_{2g} z_g + \beta_0 + e \qquad (5\text{-}2)$$

$$DF_1 = V + 2G, DF_2 = N - (V + 2G + 1)$$

Given the reference group for which $z_g = 0$, (5-2) reduces to (1). For any focal group $g \geq 1$, (5-2) reduces to (5-1). In equation (5-2), the first sum, $\sum_v b_{v0} X_v$ is the regression coefficients multiplied by the predictors in the reference group. The second sum contains g terms, and each term is an interaction between z_g and $\sum_v X_v$, the total predictor for one observation, which can be viewed as a single predictor. The dummy variable z_g is an indicator to show which of the g predictors is active given the group membership of the observation. The regression weight, β_g^*, is an estimate of the additive constant, which relates regression weights in the reference group to those in focal group g. The last term $\sum_{g=1}^{g=G} \beta_{2g} z_g$ is also a sum over g focal groups. In this term, each regression weight β_{2g} is an estimate of the difference between the intercept in the reference group (β_0) and the intercept in the focal group g.

If one assumes constant intercepts for the groups, one can just remove the varying intercepts term, $\sum_{g=1}^{g=G} \beta_{2g} z_g$, from (5-2). Below (Equation 6) is the Mean Moderated Model with an Equal intercept (MMME).

$$Y = \sum_v \beta_{v0} X_v + \sum_{g=1}^{g=G} \beta_g^* z_g \left(\sum_v X_v\right) + \beta_0 + e \qquad (6)$$

$$DF_1 = V + G, DF_2 = N - (V + G + 1)$$

The estimate of β_g^* is our estimate of the difference between the mean regression weight in the reference group and the mean regression weight in focal group g. Note that a one-point increase in the predictor variable total score leads to a change in the expected value of Y that is β_g^* units larger or smaller than the expected change in Y from a corresponding one-unit increase in the reference group. Therefore, when this model is statistically sufficient, groups differ in the relationship between the predictors and the criterion. These differences in regression weights can be written as β_g^*, which can be easily interpreted as the difference in the expected change in Y between the reference and focal group g for a corresponding one-unit change in the predictor total value.

The Mean Moderated Model also has several useful properties. The rank order of each population's regression weights (v=1, 2, …, V) remains the same in the reference group and all focal groups. The difference in regression weights between the reference group and focal group g is the same for any variable v (v=1, 2, … , V) since it is β_g^* for all V predictors. The rank order of regression weights stays the same for any group g, as in the rank order of the reference group. In other words, MMM does not change the order for the regression weights of predictors in the model.

When comparing the base model and MMM, one can test the null hypothesis, $H_0\colon \beta_g^* = 0$ for all (g ≥1) while the alternative hypothesis would be $H_a\colon \beta_g^* \neq 0$ for at least one group g ≥1. If the *F*-test is significant, that means the fuller model, here MMM, is significantly better than the base model. In other words, one size does not fit all groups and groups differ in relationships between predictors and a criterion by additive constant. When the *F*-test is not significant, however, the base model is enough and one size fits all.

When comparing MMM and FMM, one can test the null hypothesis, $H_0\colon \beta_{vg} = \beta_g^*$ for all (g ≥1), while the alternative hypothesis would be $H_a\colon \beta_{vg} \neq \beta_g^*$ for some (*v*, g ≥1) where β_{vg} is the difference in vectors of regression weights between the focal group g and the reference group. This tests whether MMM is too restrictive and FMM is needed. If the *F*-test is not significant, MMM is more parsimonious than FMM. In this case, FMM has some unnecessary parameters being estimated, and MMM is enough to explain differences between groups in relationships between predictors and a criterion. If the *F*-test is significant, MMM is too restricted and FMM is statistically better than MMM.

For example, the additive constant for Black students (β_{Black}^*) can be interpreted as the difference in expected change in Y (a criterion, mathematics achievement) between the White group (the reference group) and the Black group (focal group g) for a corresponding one-unit change in the predictor total value (total number of credits in courses completed). Therefore, the estimate of β^*_{Black} indicates that a one-point increase in the predictor total score increases the expected mathematics achievement score by β^*_{Black} more (or less) in the Black group than the White group.

When there are the same degrees of freedom for different models, it is not possible to compare them with a formal hypothesis test. For example, one cannot statistically compare PMM to MMM since they both have the same degrees of freedom ($DF_1 = V + 2G, DF_2 = N - (V + 2G + 1)$). However, one can just choose the better model by comparing one of the model fit values (R^2_{adj}, AIC, and BIC), since these models share the same degrees of freedom. In other words, one only needs to compare any of the values for model fit, since all four measures would favor the same model.

e) Linearly Moderated Models (LMM)

Both PMM and MMM models can be combined into the Linearly Moderated Model (LMM), and this model also gives each group different parameters in one equation. It allows the regression weight vectors to differ by both additive and multiplicative constants. The difference in regression weight vectors between the focal group and the reference group (β_{vg}) is $k_g\beta_{v0} + \beta^*_g$, where k_g is the constant of proportionality for the focal group g, β_{v0} is the regression weight vector in the reference group, and β^*_g is the additive constant for the focal group g.

$$Y = \sum_v \beta_{v0} X_v + \sum_{g=1}^{g=G} (k_g \beta_{v0} + \beta_g^*) z_g (\sum_v X_v) + \sum_{g=1}^{g=G} \beta_{2g} z_g + \beta_0 + e \quad (7)$$

$$DF_1 = V + 3G, DF_2 = N - (V + 3G + 1)$$

One can also express this model with a constant intercept instead of allowing intercepts to vary, which is LMME (LMM with an Equal intercept).

$$Y = \sum_v \beta_{v0} X_v + \sum_{g=1}^{g=G} (k_g \beta_{v0} + \beta_g^*) z_g (\sum_v X_v) + \beta_0 + e \quad (8)$$

$$DF_1 = V + 2G, DF_2 = N - (V + 2G + 1)$$

The constraints on these differences in vectors of regression coefficients are as a form of a linear equation using the vector of regression coefficients for the reference group. The Linearly Moderated Models are more complex than all other models suggested above except for the full moderated model. When coefficients for the two groups differ by both a multiplicative constant (k_g) and an additive constant (β_g^*), the regression weight vectors for the two groups will differ in their linear equation.

When all k_gs equal 0, LMM becomes MMM, where slopes between the reference group and focal group g differ by the additive constant, β_g^*. When all β_g^*s equal 0, LMM becomes PMM. When k_g is -1, the regression weights for that specific focal group g is equal to a constant β_g^*. When k_g equals 1, the slope for the focal group g is $2\beta_{v0} + \beta_g^*$ where the slope for the reference group is β_{v0}. Previously, PMM assumes that the rank order of the absolute value of each population's regression weights remains the same, and MMM assumes that the rank order of the value for regression coefficients does not change. Although LMM combines PMM and MMM, however, the rank order of the absolute value of each population's regression weights (v=1, 2, …, V) need not remain the same in all subgroups.

When comparing the base model and LMM, one can test the null hypothesis, H_0: $\beta_{vg} = 0$ for all (v, $g \geq 1$) while the alternative hypothesis would be H_a: $\beta_{vg} = k_g\beta_{v0} + \beta_g^* \neq 0$ for some (v, $g \geq 1$). If the F-test is significant, one size does not fit all and it is useful having differences in slopes which are linear ($\beta_{vg} = k_g\beta_{v0} + \beta_g^*$). If the F-test is not significant, however, the base model is more parsimonious than LMM, which means one size fits all.

When comparing LMM and FMM, one can test the null hypothesis, H_0: $\beta_{vg} = k_g\beta_{v0} + \beta_g^*$ for all (v, $g \geq 1$), while the alternative hypothesis would be H_a: $\beta_{vg} \neq k_g\beta_{v0} + \beta_g^*$ for some (v, $g \geq 1$). This test allows for testing whether LMM is too restrictive and we need the full moderated model. If the F-test is not significant, LMM is more parsimonious than FMM, which means FMM has some unnecessary parameters being estimated, and both proportional and additive values can represent the differences in all other regression weights between groups. If the F-test is significant, it means LMM is too restricted and the more complex model, here FMM, is significantly better than LMM.

Both PMM and MMM are nested in LMM, so when comparing PMM and LMM, one tests the null hypothesis, H_0: $\beta_g^* = 0$ for all (g ≥1), while the alternative hypothesis would be H_a: $\beta_g^* \neq 0$ at least one g ≥1. If the F-test is not significant when comparing PMM with LMM, LMM has some unnecessary parameters being estimated, and a proportional value can represent the differences in all other regression weights between groups. If the F-test is significant, LMM is needed, which means the differences in slopes differ by additive constants as well as multiplicative constants. When comparing MMM and LMM, one tests the null hypothesis, H_0: $k_g = 0$ for all (g ≥1), while the alternative hypothesis would be H_a: $k_g \neq 0$ for at least one g ≥1. If the F-test is not significant, MMM is more parsimonious, and additive values can represent the differences in all other regression weights between groups. If the F-test is significant, LMM is needed, which means the differences in slopes differ by multiplicative constants as well as additive constants.

f) ANCOVA Model

When the intercepts in the base model are allowed to vary, this leads to the familiar ANCOVA (analysis of covariance) model. This model assumes the same regression slopes for all groups but allows for different intercepts. Here, β_{v0} is the regression weight for variable v common to all groups, $\sum_{g=1}^{g=G} \beta_{2g} z_g$ gives the differences in intercepts between each focal group g and the reference group, and e is an error term.

$$Y = \sum_v \beta_{v0} X_v + \sum_{g=1}^{g=G} \beta_{2g} z_g + \beta_0 + e \tag{9}$$

$$DF_1 = V + G, DF_2 = N - (G + V + 1)$$

When comparing the base model and the ANCOVA model, one can test the null hypothesis that regression intercepts for every focal group are equal to that in the reference group. It tests the null hypothesis, H_0: $\beta_{2g} = 0$ for all (g ≥1) while H_a: $\beta_{2g} \neq 0$ for some (g ≥1). If the F-test is statistically significant, the fuller model, here the ANCOVA model is significantly better than the base model, which means intercepts differ between groups. If the F-test is not significant, the simpler, base model is more parsimonious than the ANCOVA model.

One can also compare ANCOVA with PMMU, MMMU, LMMU, and FMMU where U denotes unequal (varying) intercepts. In these model comparisons, the null and alternative hypotheses are the same when comparing these models with the base model in the prior examples. Here, hypotheses for comparing vectors of regression weights were tested without considering intercepts. For example, when comparing ANCOVA with PMMU, one can test the null hypothesis that regression weight vectors for every focal group are equal to that in the reference group. It tests the null hypothesis,

H_0: $\beta_{vg} = k_g \beta_{v0}$ for all (v, g ≥1) and $k_g \neq 0$ for some groups g ≥1, while the alternative hypothesis would be H_a: $\beta_{vg} = 0$ for all (v, g ≥1). Here, *v* denotes the regression weights for predictors and this *F*-test allows one to test whether the proportionally different regression weights exist across groups. If this *F*-test is significant, that means the fuller model, here PMM, is significantly better than the ANCOVA model. In other words, one size does not fit all groups and groups differ in relationships between predictors and a criterion. If the *F*-test is not significant, however, the simpler model is more parsimonious, and PMM has unnecessary parameters for group differences in slopes being estimated.

g) Full Moderated Model with Unequal intercept (FMMU)

The Full Moderated Model with Unequal intercept (FMMU) has varying slopes, but intercepts are constant for all groups. Relative to the full moderated model, FMMU only differs in that it has constrained intercepts. For example, when everyone has the same starting point, such as the same exam score in mathematics, but different slopes between subgroups, we may use FMMU to model the situation.

$$Y = \sum_v \beta_{v0} X_v + \sum_{g=1}^{g=G} \sum_v \beta_{vg} z_g X_v + \beta_0 + e \qquad (10)$$

$$DF_1 = (G+1)V, DF_2 = N - ((G+1)(V+1))$$

This FMMU models the situation in which groups differ in relationships, but they have a constant intercept. When using the example as for PMM, the criterion is mathematics achievement and the five predictors are credits students have completed in Algebra 2, Geometry, Statistics, Pre-Calculus, and Calculus. The slopes differ by groups but these differences in slope are not modeled by the preceding models: PMM, MMM, and LMM. Therefore, although the intercepts are equal across groups, the relationships between predictors (Algebra 2, Geometry, Statistics, Pre-Calculus, and Calculus) and mathematics achievement differ between ethnic groups, but these differences in regression weights need not be systematically modeled using either additive and/or multiplicative constants.

iii) *Level* and *Pattern*

Davison & Davenport (2002) and Davison, Davenport, Chang, Vue, & Su (2015) showed that all the predictive information contained in a group of predictors ($v \geq 2$) can be reparametrized into two variables: *Level* of the predictors and *Pattern* of the predictors. This reparameterization relies on a decomposition of both the predictor score vector and the regression weight vector into the two components (Davison et al., 2015). They define a profile as a vector of predictor scores for one observation and they decompose the profile vector into two components. The first is the *profile level* which can be indexed by the mean score $\bar{X}_p = \frac{1}{V}\sum_v X_{pv}$ or the total score $\mathrm{X}_p = \sum_v X_{pv}$ for one observation (here, a student) p. The second component is the *profile pattern*, a vector of score deviations about the profile mean: $\widetilde{\boldsymbol{X}} = \{X_v - \bar{X}\}$. They have also decomposed the regression weight vector into two components: the *level* $\bar{b} = \frac{1}{V}\sum_v b_v$ and *pattern* $\tilde{b} = \{b_v - \bar{b}\}$. Using this decomposition, the linear regression equation can be written as:

$Y = \bar{b}\sum_v X_v + b_1 \sum_v (b_v - \bar{b})(X_v - \bar{X}) + a + \varepsilon$ (Davison et al., 2017).

The first term, $\bar{b}\sum_v X_v$ is a product of the total score predictor variable $\sum_v X_v$ and the average regression weight $\bar{b}$. The corrected sum of the cross product, $\sum_v (b_v - \bar{b})(X_v - \bar{X})$, is also called *Pattern*, and it is similar to the formula for calculating covariance (Davison & Davenport, 2002). b_1 is the regression weight when *Pattern* is a predictor. Davenport et al. (2013) reported that the amount of coursework uniquely added a little (0.4%) to the variance in senior achievement while the content added additional 35% above and beyond the amount of coursework. This implies taking some courses work better than taking other courses, and therefore, the pattern of course-taking helps us to relate the content of course to mathematics achievement.

For instance, this application would be appropriate when the five predictors are credits in courses (Algebra 2, Geometry, Statistics, Pre-Calculus, and Calculus) students have completed in mathematics, mathematics achievement is the criterion, ethnic group is a moderator (z_g), and White students are our reference group as in the prior example. A student's distribution of predictors (credits completed in Algebra 2, Geometry, Statistics, Pre-Calculus, and Calculus) can be broken down into two components: *Level* (the total or mean number of completed credits in these five mathematics courses) and *Pattern* (the covariance-like statistic, having a distribution of course-taking related to high mathematics achievement) of the courses.

Using *Level* and *Pattern* makes models simpler and easier to understand while having the same amount of variance accounted for as models with many predictors. The amount of course-taking, *Level*, refers to the (mean) number of credits completed across all five courses and carries the purely quantitative information in the profile. *Pattern* is operationalized as the covariance of the credits successfully completed in each course with the regression weights from regressing the completed credits in the five courses on mathematics achievement (Davison & Davenport, 2002). *Pattern* carries the qualitative information about the content differences between the courses. The regression weights are optimal as they relate number of courses taken in each course sequence to high mathematics achievement scores (the criterion). The *Pattern* variable is computed for each student and represents the degree to which a student's course-taking over the five courses matches an optimal pattern. The optimal pattern differentiates course-taking over the five courses to distinguish high achieving students from low achieving students (Davison & Davenport, 2002). In short, *Level* contains purely quantitative information about the amount of course-taking, and *Pattern* contains qualitative information about the relationship of course-taking to achievement, and therefore these two predictors explain both quantitative and qualitative features of mathematics course-taking.

All applications are the same for the ten models except now we have only two predictors, *Level* and *Pattern*, instead of *v* predictors. For example, both effects of *Level* and *Pattern* vary across groups in the PMM models. Within a specific group, however, the change in *Level* (the number of courses) effect and *Pattern* (the contents of courses) effect are constrained by the same constant, k_g, for each focal group g. When MMM holds, the *Level* for the Black group is a β^*_{Black}-point higher than the *Level* for the White group, but *Pattern* for the Black group is not different from the *Pattern* for the White group since the additive constant does not change *Pattern*. Thus, in the MMM model, only the effect of *Level* varies between different ethnic groups, but the effect of *Pattern* is constrained equal across groups. In the LMM model, effects of both *Level* and *Pattern* are allowed to vary independently. In addition, *Pattern* is only related to the multiplicative constant, and *Pattern* in LMM is equivalent to *Pattern* in PMM when having the same multiplicative constant (k_g) regardless of the additive constant (β^*_g) and when the regression weights for the reference group (β_{v0}) are the same. If *Level* does not change over groups, but *Pattern* does, then LMM should hold, but PMM or MMM should not.

B. Substantive Review: course-taking and achievement in mathematics

Mathematics course-taking has been known to predict future academic and career success, and it is noted that there are disparities in mathematics achievement between different ethnicities (Davenport et al., 1998; Long, Conger, & Iatarola, 2012; Riegle-Crumb & Grodsky, 2010; Teitelbaum, 2003). Many studies have focused on relationships between predictors (e.g. school engagement, educational policy, course-taking) and mathematics achievement to discover the way to close the achievement gap between different groups (e.g., ethnicity, gender, English language learners). There are various methods used to quantify and examine mathematics course-taking among high school students (Adelman, 1999, 2004; Burkam & Lee, 2003; Davenport et al., 1998; Long et al., 2012; Riegle-Crumb & Grodsky, 2010; Schiller & Muller, 2003; Smith, 1996; Teitelbaum, 2003).

One method of quantifying mathematics course-taking is using the highest level of mathematics courses students had completed. Burkam and Lee (2003) used the Classification Scheme of Secondary School Courses (CSSC) to identify all 47 courses that were considered mathematics courses and noticed that most curricula followed a sequence of courses that increased in difficulty. They used an index of eight levels ranging from “No mathematics” (coded as 1) to “Advanced 3” (coded as 8), and named this the “pipeline measures.” Teitelbaum (2003) also used the highest level of mathematics courses taken when students stopped enrolling in mathematics, which condensed Burkam and Lee's (1997) eight levels into four levels of mathematics courses: Low, Medium 1, Medium 2, and Advanced. The assumption of this method is that students who complete the advanced level in mathematics courses have already mastered the skills necessary to complete all other coursework that precedes the advanced coursework.

The second approach to index mathematics course-taking is identifying the specific course(s) when students are in a certain grade (Long et al., 2012; Riegle-Crumb & Grodsky, 2010; Smith, 1996). Smith (1996) used taking Algebra prior to high school (in 8^{th} grade) to examine the effect of early entry into the advanced mathematics pipeline on mathematics achievement by the end of high school. Riegle-Crumb and Grodsky (2010) indexed taking advanced courses (beyond Algebra 2) when students were in their senior year, and Long et al. (2012) used taking a rigorous course (Level-3 math) in 9^{th} or 10^{th} grade.

Some researchers tried to make use of more than the highest course taken or enrolling in a specific course. Schiller and Muller (2003) used the total number of higher level mathematics course credits, and those higher level courses are geometry and above. Adelman (1999) tried to include both the credits of courses and types of courses taken, and he proposed the academic intensity and the quality of one's high school curriculum should be analyzed including the highest mathematics courses taken, AP courses, and remedial courses. Course-taking patterns have also been used to quantify mathematics course-taking (Davenport et al., 1998; Davenport, Davison, Wu, Kim, Kuang, Kwak, Chan, and Ayodele, 2013; Davenport, Park, Ayodele, Davison, & Kang, 2015).

Studies found that ethnic groups differed little in amount of high school coursework in mathematics, but which mathematics courses taken differed between ethnic groups (Davenport et al., 1998, 2013, 2015; Riegle-Crumb & Grodsky, 2010). Davenport et al. (2013) showed that ethnic groups had significantly lower means in mathematics achievement than that for the White group. In addition, content of mathematics coursework was more strongly associated with mathematics achievement gains than the amount of coursework, while content of coursework for both the Black and the Hispanic groups were less related to mathematics achievement compared to the White group. Riegle-Crumb and Grodsky (2010) used ethnic group as a moderator and compared students separately in advanced and nonadvanced math stratum. They focused specifically on minorities in advanced classes and found that Hispanic and African American youth from segregated schools were the most unsuccessful in terms of closing the achievement gap with their white peers (Riegle-Crumb & Grodsky, 2010).

C. Synopsis

So far, we have reviewed ten models ranging from a one size fits all to a model with varying slopes and constants by group. The base model has a constant slope and a constant intercept for all groups. Allowing intercepts to vary from the base model results in the ANCOVA model. The full moderated regression model allows both slopes and intercepts to vary across groups. Having a constant intercept relative to the full moderated model, resulted in the revised moderated model. Three separate intervening models were introduced, which are PMM, MMM, and LMM. Cook and Weisberg (2004) suggest that slopes vary across groups but only by a constant of proportionality, $k_g\beta$. Davison, Davenport, and Kohli (2017) suggest a model in which the slopes for focal group g and the reference group differ by an additive constant, β_g^*. The slopes of the linearly moderated model differ by both a multiplicative and an additive constant, $k_g\beta + \beta_g^*$. These three different intervening models with different options for slopes also have two types of intercepts, varying or constant. Therefore, we have a total of ten models with which to analyze the relationship of predictors to the criterion between multiple groups or subpopulations.

One can utilize these ten models to test which model works the best when we have different situations (slopes and/or intercepts) between different subgroups in relation to the criterion. In this thesis, specifically, I will examine ethnic group differences in relationship to mathematics course-taking and mathematics achievement using HSLS:09 data. After utilizing the ten distinct models, I will compare their results and choose the best model for this data. Each of these models can be tested against the full moderated model and/or the base model, and most can be tested against each other as well. Utilizing

Level and *Pattern* of course-taking, these models can also tell us something about the differential effects of qualitative and quantitative aspects of course-taking.

III. Methodology

A. Data

Data for this study come from the High School Longitudinal Study of 2009 (HSLS:09, Ingels et al., 2011), which is the fifth nationally representative, secondary school longitudinal study from the National Center for Education Statistics (NCES). This work builds on four studies completed to date: the National Longitudinal Study of the High School Class of 1972 (NLS:72), the High School and Beyond Longitudinal Study of 1980 (HS&B), the National Education Longitudinal Study of 1988 (NELS:88), and the Education Longitudinal Study of 2002 (ELS:2002). The base-year study for HSLS:09 involved a nationally representative sample of 944 high schools, including both public and private schools in the United States in 2009. An average of 28 students was selected based on race/ethnicity within each of the 944 participating schools via stratified two-stage sample design in the fall of 2009 when they were in the ninth grade (Ingels et al., 2011, p.81). The first follow-up took place in the spring of 2012 when most students were in the eleventh grade, since it has often been observed that students in the spring of their senior year are disengaging from high school and are not highly motivated to complete low-stakes assessments and questionnaires (Ingels et al., 2014). A postsecondary status update (2013 Update) took place in the summer and fall of 2013, and high school transcripts were collected in the 2013–14 academic year (Ingels et al., 2015). Since six schools had closed by the time of the transcript collection for the 944 base-year schools, 938 base-year schools and 3,305 transfer schools were contacted to submit student transcripts. A total of 3,028 out of 4,249 schools submitted transcripts, including 910 base-year schools and 2,118 transfer schools (Ingels et al., 2015, p.60). Therefore, a total of 21,928 transcripts were received out of 25,167 transcripts requested (Ingels et al., 2015, p.63). In this study, 21,389 students with valid transcripts having mathematics course credits were used to quantify patterns of mathematics course-taking.

B. Measures

i) Independent variables

Course catalogs for base-year schools were keyed and coded first, followed by the keying of transcripts for each of those schools (Ingels et al., 2015). The keying and coding system, a web-based data entry application, was used for both course catalog and transcript keying and coding by the keyer/coders. Cases with duplicate records in the high school transcript file were deleted. Then, only courses taken in mathematics were selected, which means only courses having a School Code for the Exchange of Data (SCED) from 02001 to 02999. This produced 142,700[*] records of courses taken by students. Courses completed with grade D- or higher or passed (n=128,060[*]) were selected; courses marked as unsatisfactory, withdraw, in progress, incomplete, and non-

[*] Unweighted sample size is rounded to the nearest ten

graded were excluded. Since the data file for high school courses and transcripts is organized by each course, credits earned for courses having the same SCED codes were summed to make course variables for each student. For instance, the 'sum02001' variable is calculated by summing all credits completed in courses in this category (having the SCED code of '02001') for each student. Therefore, variables representing credits completed in each of 67 SCED mathematics courses were produced. Data for all students in this study include 67 separate variables, and these variables are used to quantify course-taking in mathematics and to predict mathematics achievement in the Spring of 2012 (the number of students, *n*=21,410[*]). Thirteen sequences based on the 67 SCED mathematics courses were generated using the logic of the highest level of mathematics course taken (X3THIMATH). X3THIMATH has a total of 14 categories from No math to AP/AB Calculus (0: No math, 1: Basic Math, 2: Other Math, 3: Pre-Algebra, 4: Algebra I, 5: Geometry, 6: Algebra II, 7: Trigonometry, 8: Other advanced math, 9: Probability and statistics, 10: Other AP/IB math, 11: Precalculus, 12: Calculus, 13: AP/IB Calculus), so thirteen mutually exclusive sequences using the sum of credits completed in those sequences of courses from Basic Math to AP/IB Calculus were made. For instance, to make the variable, Probability and statistics, the credits completed in the courses SCED 02201 (Probability and Statistics), 02202 (Inferential Probability and Statistics), 02204 (Particular Topics in Probability and Statistics), 02207 (Probability and

Statistics—Independent Study), and 02209 (Probability and Statistics—Other) were summed together.

The two independent variables in this study, *Level* and *Pattern*, were reproduced using the 13 sequences of mathematics course-taking. Here *Level* is the mean of credits completed in the 13 sequences ($\frac{1}{V}\sum_v X_{pv}$) and *Pattern* is the corrected sum of the cross product, $\sum_v (b_v - \bar{b})(X_v - \bar{X})$, similar to the equation for covariance when divided by *n*. Therefore, separate analyses were performed to differentiate between the two types of independent variables: 13 sequences of predictors or *Level* and *Pattern.*

ii) Moderator variable

A moderator variable, ethnicity, is recorded from X2RACE having eight categories (1: American Indian/Alaska Native, non-Hispanic, 2: Asian, non-Hispanic, 3: Black/African-American, non-Hispanic, 4: Hispanic, no race specified, 5: Hispanic, race specified, 6: More than one race, non-Hispanic, 7: Native Hawaiian/Pacific Islander, non-Hispanic, and 8: White, non-Hispanic). A total of 2,110* students in More than one race (X2RACE=6, *n*=1,850*, 8.7%), American Indian/Alaska Native, non-Hispanic (X2RACE=1, *n*=150*, .7%) and Native Hawaiian/Pacific Islander, non-Hispanic (X2RACE=7, *n*=100*, .5%) categories were not included in this study, resulting in a final sample of 19,310* students. The new 'ethnicity' variable is a moderator, which is composed of four categories of ethnic groups (*n*=19,310*), which are Asian (X2RACE=2), Black/African American (X2RACE=3), Hispanic (X2RACE=4 or 5), and White (X2RACE=8). There are 19,310* students whose records included valid course-taking information and 'ethnicity' corresponding to one of the four groups.

iii) Dependent variable

The dependent variable of mathematics achievement, X2TXMSCR, which is the math IRT-estimated number correct, is an overall criterion-referenced measure at the time of the first follow-up assessment. The estimated number-correct score is an estimate of the number of items that students would have answered correctly had they responded to all items in the respective item pool defined for each criterion, based on the pattern of correct answers (Ingels et al., 2014). The term criterion in "criterion-referenced measure" means the set of skills defined by the assessment framework and represented by the assessment item pool. The estimated number-correct score (X2TXMSCR) has a potential range of 0 to 118 since the sum of the probabilities across all 118 items is the estimate for the number of items the student would have gotten correct.

iv) Covariates

In addition, the mathematics score at the time of the base-year assessment in mathematics, X2X1TXMSCR, was also used as a covariate indicating prior mathematics achievement. In the base year, the IRT-estimated number correct is about 72 items, but the item pool increased because of new items being fielded in the first follow-up. Accordingly, the base-year number-correct scores were re-estimated to reflect performance on the entire item pool of 118 unique items using a student's base-year theta (ability score). These scores are recommended to be used as longitudinal measures of overall growth, defined as a measure of change in the aggregate with the availability of assessment scores at two points in time (Bozick & Ingels, 2008; Ingels et al., 2014).

A total of four families of analyses were conducted, regressing mathematics achievement on four types of predictors using 1) 13 sequences of course-taking, 2) *Level* and *Pattern*, 3) 13 sequences of course-taking and prior achievement, and 4) *Level* and *Pattern* with prior achievement. Predicting math achievement using 13 sequences of course-taking comes first. Analysis using both *Level* and *Pattern* instead of the 13 sequences comes next to show the similarities and differences between models with or without prior achievement. After testing possible covariates (prior achievement and SES), analyses with prior achievement were performed to control the effect of prior achievement as a covariate. All four families of analyses were conducted separately to investigate using two different types of predictors and with or without prior achievement (See Figure 4) to show the relationship between course-taking and achievement.

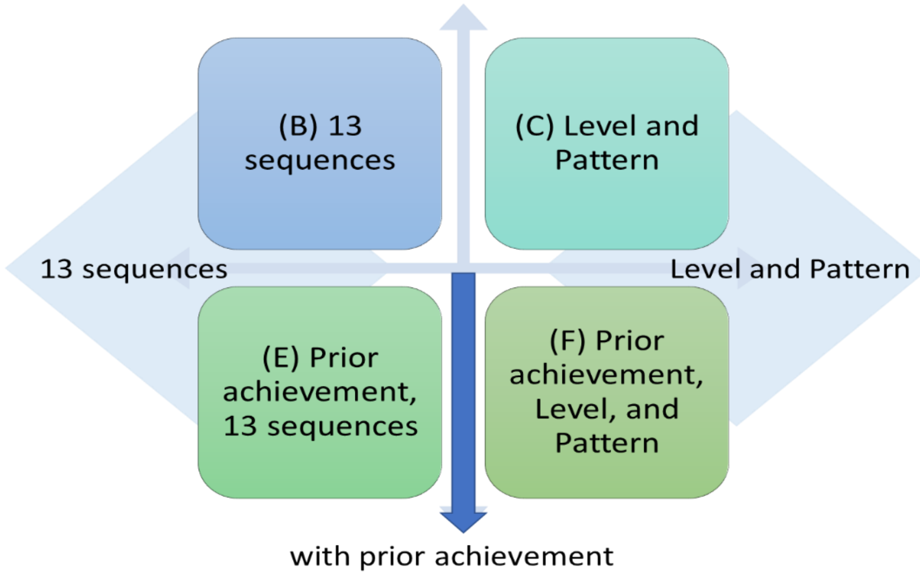


Figure 4. Four families of analyses (B, C, E and F sections in Chapter IV) with two types of predictors and with or without the presence of a covariate (prior achievement)

In addition, after performing listwise deletion, including credits completed in Mathematics and both pre- and post-achievement scores, a total of 15,750* students with valid transcripts, complete pre- and post-achievement scores, and ethnicity grouped into one of the four categories were used in this study. The table for frequencies of the ethnicity variable and the average of post achievement scores by ethnicity is given below (Table 3.1).

Table 3.1
Frequency table and average post achievement scores by ethnic group

Ethnicity	Frequency*	Percent†	Average Post Achievement Score†
White	9,880	58.54	68.887
Asian	1,400	3.88	80.636
Black/African American	1,720	14.08	57.105
Hispanic	2,740	23.50	61.664
Total	15,750	100.00	65.986

Note. *Unweighted sample size numbers are rounded to the nearest ten; †weighted using transcript weight

C. Methods

i) Ten multiple regression models

In chapter 2, an integrated review of the ten multiple regression models was introduced. These ten models are (1) the simplest base model (simple multiple regression model), (2) the most complex full moderated multiple regression model (FMM; Saunders, 1956), (3) the Proportional Moderated Model (PMM; Cook & Weisberg, 2004), (4) the Mean Moderated Model (MMM; Davison et al., 2017), (5) the Linearly Moderated Model (LMM; combination of PMM and MMM models), (6) the ANCOVA model (allowing intercepts to vary between groups relative to the simple multiple regression model), (7) the FMM with a constant intercept (FMME), (8) the PMM with a constant intercept (PMME), (9) the MMM with a constant intercept (MMME), and (10) the LMM with a constant intercept (LMME).

ii) Model comparison

George Box made the famous statement, "All models are wrong but some are useful (Box & Draper, 1987)," and Occam's razor highlights the importance of parsimony by saying, "Shave away all but what is necessary" (Burnham & Anderson, 2004). Goodness-of-fit must be balanced to avoid overfitting, since inference under models with too few parameters (variables) can be biased, while models having too many parameters (variables) may have poor precision or identification of effects that are spurious (Judd, McClelland, & Ryan, 2009).

Each of the above models will be applied to the same data with resultant model comparisons (the model and error degrees of freedom are given above for each model). Standard software is readily available for most of the models except for the four models that include the proportional model (PMM). In all instances for model comparison we use:

$$F_{(DFe(1)-DFe(2),\ DFe(2))} = \frac{SSE_{(1)} - SSE_{(2)}}{df_{e(1)} - df_{e(2)}} / \frac{SSE_{(2)}}{df_{e(2)}} \qquad (11)$$

In Equation 11, (*2*) represents the fuller model, the model with more parameters being estimated, and (*1*) represents the restricted model with fewer parameters being estimated. For example, when one compares models while allowing intercepts to vary versus not vary, the difference in R^2 means the amount of variability increased by allowing intercepts to vary. Likewise, all other models can be compared with a full moderated model since they are subsets of the full moderated model. One can also compare all models with a simple multiple regression model (one-size fits all model) since it is a subset of all other models. Note that only nested models, when it is a subset of the other model, were compared to each other in this study.

Results are given for each model comparison and interpretations are rendered based on whether the test was significant or not. In addition to statistical tests, this study uses several descriptive indices for the comparisons. These include R^2, R^2_{adj}, Akaike information criterion (*AIC*, Burnham & Anderson, 2004), and Bayesian information criterion (*BIC*, Burnham & Anderson, 2004). The model is considered better when having a higher R^2_{adj} value or lower *AIC* and *BIC* values. Multiple R^2 is the amount of variability in Y explained by predictors, defined by $(1 - SSE/SST)$. The adjusted multiple R^2 considers and corrects for the upward bias in R^2, and consequently punishes variables (including irrelevant ones) in the model ($R^2_{adj} = 1 - (1 - R^2) * (n - 1)/(n - V - 1)$, where *n* is the sample size and *V* is the number of predictors in the model. AIC is an asymptotically unbiased estimator and it represents information-theoretic selection based on Kullback-Leibler (K-L) information loss (Burnham & Anderson, 2004). Both AIC and BIC indices provide the comparison of model fit in models that are not nested as well as nested models, and they also take into account the number of regression coefficients being tested (Cohen, Cohen, West, & Aiken, 2003). Computation of the AIC is based on the likelihood under the model containing $(V+1)$ parameters (V predictors and one intercept) and where n is the number of cases and $\hat{L}$ is the maximized value of the likelihood function of the model: AIC=$-2Ln(\hat{L}) + 2(V + 1)$ and BIC=$-2Ln(\hat{L}) + Ln(n)(V + 1)$. In the special case of least squares (LS) estimation with normally distributed errors, AIC can be expressed as $n \times Ln(\sigma^2) + 2(V + 1)$ and BIC can be written as $n \times Ln(\sigma^2) + Ln(n)(V + 1)$ where n is the sample size, Ln is the natural log, the Mean Square Error is the best estimate of σ^2, and V is the number of predictors in the model. BIC approximation follows Bayesian model selection based on Bayes factors. BIC assumes that there is a true model, independent of n, that generates the data. AIC assumes the truth is unknown, however, and it tries to select the best model with minimum AIC values.

As already discussed, model selection is a goodness-of-fit versus parsimony trade-off. In other words, an excellent fit to the data (high descriptive adequacy) tends to be achieved with a highly complex model, which is related to low parsimony and low generalizability (Pitt, Myung, & Zhang, 2002). Therefore, considerations call for a balance between underfitted and overfitted models (Forster, 2000). Here, the model comparison is whether the increment in the predicted or explained variability, obtained by adding predictors, is significant. For example, when the F-test is significant comparing the base model and MMM, one size does not fit all and additive slopes are needed for different subgroups. When the F-test is not significant in a comparison of

MMM and the most complex model (FMM), FMM fits no better than MMM when considering both the model fit and the parsimony of the model.

When the degrees of freedom for both models are the same, one can refer to the final table of model fit measures (R^2, R^2_{adj}, AIC, and BIC). For instance, one cannot statistically compare PMM to MMM using an F-test, since they both have the same degrees of freedom ($DF_1 = V + 2G$, $DF_2 = N - (V + 2G + 1)$). In this case, one can choose the better model by comparing any of the model fit values. When both PMM and MMM are found to be statistically better than the simplest model and they are also more parsimonious than the most complex model, one may need to compare R^2_{adj} values directly instead of performing the F-test since their model and error degrees of freedom are the same. Therefore, the model with higher R^2_{adj} value fits better than the other one in this case.

iii) Statistical program

IBM SPSS Statistics 24 was used for cleaning and handling the data and the R program (Version 3.4.2) was used to analyze the ten multiple regression models. Specifically, I used linear model function (lm) in R, which uses least squares to estimate parameters and errors. Furthermore, my program incorporates analytic weights to account for over and under sampling and missing data. W3HSTRANS, the weight specific to high school transcript response, was used in this study. Listed for each of the ten models is the parameter estimates with standard errors, t-values, and significance levels for the t-statistic.

Models having the same slope or varying slopes, including the base model, the ANCOVA model, FMME, and FMMU, do not need special treatment. However, a special program is needed to analyze PMM, MMM, and LMM. For example, in order to analyze PMM, a new variable (e.g. NC) needs to be calculated using a subset of participants: those who are in the reference group (ethnicity=0). NC ($\sum_v \beta_{v0} X_{v0}$) is obtained by sum of the product of regression weights from the reference group and all v predictors for the reference group where v means the number of predictors and 0 means participants in the reference group (ethnicity=0). When adding g predictors to the base model for PMME (or to the ANCOVA model for PMMU), one for each focal group g, each β_{vg} is calculated using NC multiplied by one constant (k_g) for each focal group g, $\beta_{vg} = k_g \times \beta_{v0}$. For MMM, sum of a total predictor values (e.g. Sum) needs to be added as a new variable. For LMM both the g-number of 'NC' variables and 'Sum' variables need to be added to the base model or the ANCOVA model depending on intercepts. When allowing intercepts to vary (models ending with U), we need to add a dummy variable for each focal group g. Here, there are three focal groups except for the reference group (White).

Final two tables (4.X.11 and 4.X.12, where X is B, C, E, and F) summarized 1) goodness of fit measures including R^2, R^2_{adj}, AIC, and BIC and 2) F-test results for nested models. The R^2 and R_{adj}^2 are given using the 'summary()' command in the R program, but R_{adj}^2 needs to be recalculated using a new mean squared error accounting for the design effect. AIC and BIC are also calculated using a new mean squared error accounting for the design effect. Secondly, *F*-test results between nested models are also produced. An F-value, the difference in degrees of freedom between the restricted model and the fuller model, the degrees of freedom for the fuller model, and corresponding p-values are produced for each comparison.

iv) Design effect

The design effect (deff) is the ratio of the actual variance, under the sampling method used, to the variance calculated under the assumption of simple random sampling (Sturgis, 2004). HSLS: 09 data used a stratified, two-stage sample design with primary sampling units (schools) selected in the first stage, and students randomly selected from the sampled schools in the second stage (Ingels et al., 2011).

For HSLS: 09 data, the average design effects for students' ethnicities are 4.0, 4.9, 3.7, 2.7, and 3.1 for Hispanic, Asian, Black, White, and more than one race, respectively (Ingels et al., 2014. p.126), so 4.0 was used as the design effect for this study because it is the largest effect for all groups except the Asian group. It is also larger and more conservative than the mean of the average design effects for each ethnic group (3.68). To employ the design effect of 4, the mean square error (MSE) with an effective sample size (n/4) was calculated with the Microsoft Excel program. It is calculated by dividing n by 4 (design effect) instead of using n. The t-values for each parameter are recalculated using the newer MSE, and so standard errors and significance levels for the t-statistic are also adjusted for the design effect. For instance, when using the design effect of 4 in this study, the new MSE is calculated using the prior MSE multiplied by 4. Using the new MSE, we can calculate new t-values for each parameter.

IV. Results

A. Descriptive statistics and correlation

Descriptive statistics for all variables used in the analyses are presented in Table 4.A.1. The minimum of zero credit means students took no credits in a certain sequence of mathematics course-taking. The maximum credits taken are 6 in the “Other advanced” sequence and 8 in the “Other Math” sequence, each of which is greater than the average credits most high school students take in Mathematics. This is because some students took a total of 8 credits, all of them in the “Other Math” sequence or a total of 6 credits, all of them in the “Other advanced” sequence in high school. The maximum value in Basic Math is 4.04, and this is because some schools have trimester systems and they put .34 instead of .33 for one credit per year for one trimester. *Level* is the mean of credits taken in all sequences of courses. The average post achievement score is higher than the average prior achievement score, and post achievement varies more than prior achievement.

Table 4.A.1

Descriptive Statistics with transcript weights for overall sample (n=15,750[*])

Variables	Min	Max	Mean	SD
Basic Math	0.00	4.04	0.15	0.42
Other Math	0.00	8.00	0.17	0.51
Pre-Algebra	0.00	3.00	0.11	0.33
Algebra I	0.00	5.50	0.81	0.59
Geometry	0.00	4.25	0.88	0.45
Algebra II	0.00	4.00	0.67	0.54
Trigonometry	0.00	2.50	0.17	0.38
Other Advanced	0.00	6.00	0.23	0.62
Probability and Statistics	0.00	2.00	0.06	0.23
Other AP/IB Math	0.00	2.00	0.06	0.23
Pre-Calculus	0.00	3.00	0.30	0.46
Calculus	0.00	3.00	0.05	0.22
AP/IB Calculus	0.00	4.00	0.12	0.36
Level	0.00	1.00	0.29	0.09
Pattern	-6.18	3.06	-0.51	0.81
Prior Achievement (base year)[†]	25.09	103.79	55.43	15.57
Post Achievement (first follow-up)[†]	25.01	115.10	65.99	18.83

Note. [*]Number of participants are rounded to the nearest ten; [†]IRT-estimated number right score for 118 items

The descriptive statistics for all variables by ethnic group are presented in Table 4.A.2. The Asian group has the highest prior and post achievement scores followed by the White group, the Hispanic group, and the Black group. Credits completed in thirteen subjects vary by group, and the Asian group tends to take more higher-level courses compared to the Black and the Hispanic groups. Similarly, *Pattern* of the Asian group is the highest among all groups followed by the White group, the Black group, and the Hispanic group. Differences exist in course-taking and *Pattern* as well as prior and post achievement between ethnic group while the mean of *Level* seems to be similar except for the Asian group.

Table 4.A.2

Descriptive Statistics with transcript weights by ethnic group

Ethnic group	Variables	Min	Max	Mean	SD
White (n=9,880*)	Basic Math	0.00	4.04	0.14	0.42
	Other Math	0.00	8.00	0.16	0.49
	Pre-Algebra	0.00	3.00	0.11	0.33
	Algebra I	0.00	4.00	0.77	0.58
	Geometry	0.00	4.25	0.87	0.43
	Algebra II	0.00	3.00	0.69	0.53
	Trigonometry	0.00	2.00	0.19	0.40
	Other Advanced	0.00	6.00	0.23	0.61
	Probability and Statistics	0.00	2.00	0.07	0.24
	Other AP/IB Math	0.00	2.00	0.06	0.24
	Pre-Calculus	0.00	3.00	0.34	0.48
	Calculus	0.00	3.00	0.07	0.25
	AP/IB Calculus	0.00	3.00	0.13	0.38
	Level	0.00	0.77	0.29	0.09
	Pattern	-3.93	2.75	-0.43	0.82
	Prior Achievement	25.09	103.79	57.96	15.28
	Post Achievement	26.18	115.10	68.89	18.56
Asian (n=1,400*)	Basic Math	0.00	3.00	0.07	0.27
	Other Math	0.00	8.00	0.10	0.38
	Pre-Algebra	0.00	2.00	0.04	0.19
	Algebra I	0.00	5.50	0.65	0.70
	Geometry	0.00	3.50	0.84	0.41
	Algebra II	0.00	4.00	0.73	0.51
	Trigonometry	0.00	2.00	0.20	0.42
	Other Advanced	0.00	5.00	0.24	0.59
	Probability and Statistics	0.00	2.00	0.04	0.20
	Other AP/IB Math	0.00	2.00	0.16	0.36
	Pre-Calculus	0.00	3.00	0.53	0.51
	Calculus	0.00	3.00	0.08	0.27
	AP/IB Calculus	0.00	4.00	0.47	0.67
	Level	0.04	0.77	0.32	0.09

	Pattern	-3.93	3.06	0.15	0.99
	Prior Achievement	25.33	103.79	67.28	16.37
	Post Achievement	28.61	114.29	80.64	18.86
Black/ African American (*n*=1,720*)	Basic Math	0.00	4.00	0.18	0.46
	Other Math	0.00	7.00	0.27	0.70
	Pre-Algebra	0.00	3.00	0.17	0.41
	Algebra I	0.00	3.33	0.87	0.52
	Geometry	0.00	3.50	0.86	0.51
	Algebra II	0.00	2.50	0.61	0.55
	Trigonometry	0.00	2.50	0.18	0.39
	Other Advanced	0.00	6.00	0.33	0.82
	Probability and Statistics	0.00	1.00	0.06	0.24
	Other AP/IB Math	0.00	2.00	0.03	0.17
	Pre-Calculus	0.00	2.00	0.20	0.40
	Calculus	0.00	2.00	0.02	0.15
	AP/IB Calculus	0.00	2.00	0.04	0.21
	Level	0.00	0.62	0.29	0.09
	Pattern	-3.44	1.89	-0.78	0.68
	Prior Achievement	25.10	97.09	47.62	13.98
	Post Achievement	26.76	112.58	57.11	16.35
Hispanic (*n*=2,740*)	Basic Math	0.00	4.00	0.17	0.43
	Other Math	0.00	8.00	0.17	0.44
	Pre-Algebra	0.00	3.00	0.10	0.31
	Algebra I	0.00	4.25	0.90	0.63
	Geometry	0.00	4.00	0.91	0.45
	Algebra II	0.00	4.00	0.64	0.57
	Trigonometry	0.00	2.00	0.10	0.32
	Other Advanced	0.00	4.00	0.18	0.49
	Probability and Statistics	0.00	2.00	0.05	0.20
	Other AP/IB Math	0.00	2.00	0.05	0.22
	Pre-Calculus	0.00	2.00	0.22	0.42
	Calculus	0.00	2.00	0.03	0.16
	AP/IB Calculus	0.00	3.00	0.07	0.29
	Level	0.00	1.00	0.28	0.10
	Pattern	-6.18	2.40	-0.65	0.73
	Prior Achievement	25.32	103.79	51.86	14.16
	Post Achievement	25.01	111.14	61.66	17.59

Note. *number of participants in each group are rounded to the nearest ten.

The correlation table detailing prior achievement, post achievement, *Level*, and *Pattern* (overall and within ethnicity) to allow one to ascertain the incremental effect of prior achievement and course-taking is presented below. Comparing ethnic groups, weaker correlations between *Pattern* and post achievement for Black/African American and Hispanic groups suggest that the relationship of these Patterns of course-taking to achievement are not the same for each ethnic group. These minority groups have less optimal patterns compared to the White group, and groups having optimal patterns do better.

Table 4.A.3

Correlations among prior and post mathematics achievement, Level, and Pattern for overall and by ethnic group with transcript weights

	Group	Prior Achievement	Post Achievement	*Level*	*Pattern*
Overall	Prior Achievement	1.00			
	Post Achievement	0.74	1.00		
	Level	0.26	0.30	1.00	
	Pattern	0.56	0.58	0.07	1.00
White	Prior Achievement	1.00			
	Post Achievement	0.73	1.00		
	Level	0.29	0.33	1.00	
	Pattern	0.58	0.60	0.16	1.00
Asian	Prior Achievement	1.00			
	Post Achievement	0.79	1.00		
	Level	0.26	0.24	1.00	
	Pattern	0.63	0.64	0.16	1.00
Black/ African American	Prior Achievement	1.00			
	Post Achievement	0.66	1.00		
	Level	0.21	0.23	1.00	
	Pattern	0.34	0.39	-0.25	1.00

Hispanic	Prior Achievement	1.00			
	Post Achievement	0.68	1.00		
	Level	0.21	0.25	1.00	
	Pattern	0.47	0.47	-0.06	1.00

Note. All correlations are $p < .001$

Below are sample results for some participants to show the difference in observations based on *Level* and *Pattern* as they relate to a host of other variables. The first three students show that *Pattern* of course-taking varies although they share the same *Level* of course-taking. The fourth student took six credits in mathematics but both prior and post achievement scores are lower than other students. The fifth student took only three credits in mathematics during high school, but this student's scores for both prior and post achievement were the highest among all five students. These two students demonstrate that the content of courses taken are not always aligned with the number of courses taken. Therefore, both *Level* and *Pattern* of course-taking should be considered to account for the effect of mathematics course-taking on mathematics achievement.

Table 4.A.4.

Number of Carnegie Units in Each Sequence and Other Scores for Five Students

Course Sequence	Student1	Student2	Student3	Student4	Student5
Basic Math	-	-	-	2.00	-
Other Math	2.00	-	-	-	-
Pre-Algebra	-	1.00	-	1.00	-
Algebra I	1.00	-	-	1.00	-
Geometry	1.00	1.00	1.00	1.00	-
Algebra II	-	1.00	1.00	1.00	-

Trigonometry	-	-	-	-	2.00
Other Advanced	-	-	-	-	-
Probability and Statistics	-	-	-	-	-
Other AP/IB Math	-	-	-	-	-
Pre-Calculus	-	1.00	1.00	-	-
Calculus	-	-	1.00	-	-
AP/IB Calculus	-	-	-	-	1.00
	-	-	-	-	-
Level	0.31	0.31	0.31	0.46	0.23
Pattern	-1.66	-0.24	0.87	-2.40	1.22
Prior Achievement	40.24	57.50	72.41	32.58	84.11
Post Achievement	56.46	78.19	82.99	48.44	99.51

B. Ethnic differences in the relationship between thirteen sequences of mathematics course-taking and mathematics achievement

In this section, mathematics achievement is predicted by credits for each of thirteen sequences of mathematics course-taking (X1: Basic Math, X2: Other Math, X3: Pre-Algebra, X4: Algebra I, X5: Geometry, X6: Algebra II, X7: Trigonometry, X8: Other advanced math, X9: Probability and statistics, X10: Other AP/IB math, X11: Precalculus, X12: Calculus, and X13: AP/IB Calculus), but moderated by ethnicity with White being the referent group (0: White, 1: Asian, 2: Black/African American, and 3: Hispanic).

i) Base model

The base multiple regression model (Table 4.B.1) is the simplest model, and it assumes all regression weights are equal across ethnicity (one size fits all). For this model, we estimated one optimal intercept and one optimal slope coefficient per predictor. The results imply that, on average, taking a credit in basic mathematics courses categorized as “Basic Math,” “Pre-Algebra,” “Other Math,” and “Algebra I” were negatively related to mathematics achievement, controlling for number of credits taken in other sequences of mathematics courses (b = -3.42, -2.83, -2.21, and -1.31, respectively). Sequences from “Algebra II” to “AP/AB Calculus” are positively related to mathematics achievement. For example, taking a credit in “AP/AB Calculus” is associated with a 14.13-point increase in mathematics achievement.

Table 4.B.1

Results for the base model predicting mathematics achievement

	Estimate	SE*	t*	p*
(Intercept)	56.75	0.84	67.45	0.00
Basic Math	-3.42	0.58	-5.93	0.00
Other Math	-2.21	0.47	-4.69	0.00
Pre-Algebra	-2.83	0.73	-3.89	0.00
Algebra I	-1.31	0.43	-3.04	0.00
Geometry	0.86	0.57	1.53	0.13
Algebra II	4.00	0.48	8.39	0.00
Trigonometry	7.10	0.65	10.95	0.00
Other Advanced	1.58	0.41	3.82	0.00
Probability and Statistics	5.02	1.01	4.99	0.00
Other AP/IB Math	8.23	1.03	8.00	0.00
Pre-Calculus	11.61	0.57	20.27	0.00
Calculus	11.56	1.06	10.87	0.00
AP/IB Calculus	14.13	0.70	20.29	0.00
R^2		0.3994		
Adjusted R^2*		0.3974		
AIC*		47253.889		
BIC*		47341.783		

Note. *after accounting for weights and the design effect of 4

When plotting the regression weights for thirteen sequences of mathematics course-taking (Figure 5), one can see the rising trends supporting that highest mathematics course taken positively relates to mathematics achievement (except for ‘Other’ categories [Other Math, Other Advanced, Other AP/AB Math]).

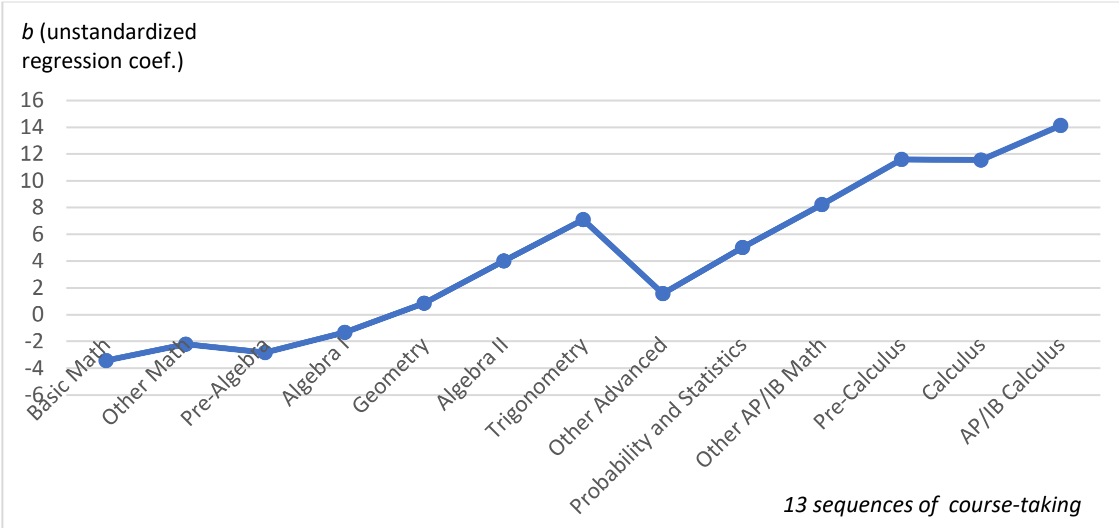


Figure 5. Unstandardized regression coefficients to predict mathematics achievement with thirteen course-taking categories for overall students

ii) ANCOVA

The ANCOVA model (Table 4.B.2) assumes the same regression slopes for all groups, but allows different intercepts. In doing so, this model tests whether the means differ between ethnic groups while controlling for the effect of other predictors. The predicted means of all focal groups statistically differ from that of the reference group of White students. Asian students had the highest predicted mathematic achievement score, on average, which was 3.29 points higher than the mean of the White group. Meanwhile, the Black/African American group and the Hispanic group had lower predicted mean mathematics achievement scores (7.52 and 3.46 points lower, respectively) than the predicted mean mathematics achievement score for White student. The order of regression coefficients is the same as the order in the base model.

Table 4.B.2

Results for the ANCOVA model predicting mathematics achievement

	Estimate	SE*	t*	p*
(Intercept)	58.37	0.85	68.57	0.00
Basic Math	-3.32	0.57	-5.86	0.00
Other Math	-1.91	0.46	-4.11	0.00
Pre-Algebra	-2.55	0.72	-3.55	0.00
Algebra I	-1.05	0.42	-2.47	0.01
Geometry	1.07	0.56	1.92	0.05
Algebra II	3.86	0.47	8.21	0.00
Trigonometry	7.00	0.64	10.93	0.00
Other Advanced	1.77	0.41	4.33	0.00
Probability and Statistics	4.88	0.99	4.93	0.00
Other AP/IB Math	7.89	1.01	7.81	0.00
Pre-Calculus	11.15	0.56	19.74	0.00
Calculus	10.79	1.05	10.32	0.00
AP/IB Calculus	13.35	0.69	19.24	0.00
Asian†	3.29	1.22	2.70	0.01
Black/African American†	-7.52	0.69	-10.94	0.00
Hispanic†	-3.46	0.57	-6.09	0.00
R^2		0.4212		
Adjusted R^2*		0.4189		
AIC*		47116.942		
BIC*		47223.670		

Note. *after accounting for weights and the design effect of 4; †the predictive mean difference (intercept) between this focal group and the reference (White) group

iii) PMM

PMM with a constant intercept (PMME, Table 4.B.3) constrains differences in slopes between the focal group and the reference group to be a multiplicative constant of the slopes without allowing intercepts to vary by group. The multiplicative constant for the Asian group was not statistically different from 0, which means the difference between vectors of all regression weights (Basic Math to AP/IB Calculus) between the Asian group and the White group is not statistically different from the vector of all regression weights for the White group of students. The multiplicative constants for the Black/African American group and the Hispanic group were statistically significant (-.41 and -.20 respectively) suggesting the vector of regression weights for the White group was multiplied by -.41 to get the difference in vectors of regression weights between the Black/African American group and the White group. The vector of regression weights for the White group was multiplied by -.20 to get the difference in vectors of regression weights (Basic Math to AP/IB Calculus) between the Hispanic group and the White group. The negative constants of proportionality lead to these two focal groups having smaller absolute values of the regression coefficients compared to the regression coefficients for the reference group.

Table 4.B.3

Results for PMME predicting mathematics achievement

	Estimate	SE*	t*	p*
(Intercept)	56.44	0.84	67.56	0.00
Basic Math	-3.83	0.58	-6.65	0.00
Other Math	-2.44	0.47	-5.21	0.00
Pre-Algebra	-3.13	0.72	-4.32	0.00
Algebra I	-1.36	0.43	-3.17	0.00
Geometry	1.29	0.56	2.28	0.02
Algebra II	4.49	0.48	9.36	0.00
Trigonometry	7.65	0.65	11.82	0.00
Other Advanced	1.92	0.41	4.64	0.00
Probability and Statistics	5.24	1.00	5.24	0.00
Other AP/IB Math	8.62	1.03	8.38	0.00
Pre-Calculus	12.38	0.58	21.41	0.00
Calculus	11.46	1.06	10.86	0.00
AP/IB Calculus	14.24	0.73	19.62	0.00
k_{Asian}	0.05	0.05	0.89	0.37
$k_{Black/African\ American}$	-0.41	0.06	-6.91	0.00
$k_{Hispanic}$	-0.20	0.04	-4.75	0.00
R^2		0.4093		
Adjusted R^2*		0.4069		
AIC*		47197.101		
BIC*		47303.829		

Note. *after accounting for weights and the design effect of 4

All multiplicative constants (k_gs) in PMM with unequal intercepts (PMMU) were not statistically significant after allowing differences in intercepts (Table 4.B.4). Allowing for intercept changes, the Asian group, on average, had higher predicted mean mathematics achievement score by 5.78 points than the White group. The Black/African American group and the Hispanic group, on average, had lower predicted scores (6.99

and 2.78 points lower, respectively) in mathematics achievement than those for the White group.

Table 4.B.4

Results for PMMU predicting mathematics achievement

	Estimate	SE*	t*	p*
(Intercept)	58.05	0.87	66.75	0.00
Basic Math	-3.43	0.57	-6.01	0.00
Other Math	-1.98	0.47	-4.23	0.00
Pre-Algebra	-2.61	0.72	-3.62	0.00
Algebra I	-1.13	0.43	-2.66	0.01
Geometry	1.13	0.56	2.02	0.04
Algebra II	4.04	0.48	8.38	0.00
Trigonometry	7.17	0.65	11.06	0.00
Other Advanced	1.83	0.41	4.46	0.00
Probability and Statistics	5.00	0.99	5.05	0.00
Other AP/IB Math	8.21	1.02	8.05	0.00
Pre-Calculus	11.45	0.59	19.47	0.00
Calculus	10.94	1.05	10.45	0.00
AP/IB Calculus	13.83	0.73	18.97	0.00
Asian†	5.78	2.02	2.86	0.00
Black/African American†	-6.99	0.85	-8.26	0.00
Hispanic†	-2.78	0.70	-3.99	0.00
k_{Asian}	-0.14	0.09	-1.64	0.10
$k_{Black/African\ American}$	-0.06	0.07	-0.86	0.39
$k_{Hispanic}$	-0.08	0.05	-1.59	0.11
R^2		0.4219		
Adjusted R^2*		0.4191		
AIC*		47121.056		
BIC*		47246.618		

Note. *after accounting for weights and the design effect of 4; †the predictive mean difference (intercept) between this focal group and the reference (White) group

iv) MMM

MMM constrains the differences in slopes between the focal group and the reference group to be a constant. When all regression coefficients between a focal group g and a reference group differ by an additive constant, β_g^*, the vectors of regression coefficients for the two groups will differ only in their means. For MMM with a constant intercept (MMME, Table 4.B.5), the negative values of $\beta_{Black/African\ American}^*$ and $\beta_{Hispanic}^*$ indicate that lower regression coefficients for the Black/African American group (-1.74) and Hispanic group (-.87), in which a one-point increase in total predictor score leads to a 1.74 point and a 0.87 point decrease in the expected mathematics achievement score compared to the expected mathematics achievement score for the White group. In other words, increasing total predictor score leads to expected loss in mathematics achievement for the Black/African American group and the Hispanic group by 1.74 point and 0.87 point, respectively, when intercepts are constrained to be the same.

Table 4.B.5

Results for MMME predicting mathematics achievement

	Estimate	SE*	t*	p*
(Intercept)	56.29	0.83	67.82	0.00
Basic Math	-2.77	0.57	-4.85	0.00
Other Math	-1.39	0.47	-2.95	0.00
Pre-Algebra	-2.05	0.72	-2.84	0.00
Algebra I	-0.62	0.43	-1.43	0.15

Geometry	1.64	0.56	2.92	0.00
Algebra II	4.43	0.47	9.38	0.00
Trigonometry	7.44	0.64	11.60	0.00
Other Advanced	2.29	0.41	5.54	0.00
Probability and Statistics	5.26	0.99	5.30	0.00
Other AP/IB Math	8.29	1.02	8.16	0.00
Pre-Calculus	11.64	0.56	20.62	0.00
Calculus	11.15	1.05	10.64	0.00
AP/IB Calculus	13.79	0.70	19.70	0.00
β^*_{Asian}	0.54	0.29	1.90	0.06
$\beta^*_{Black/African\ American}$	-1.74	0.17	-10.06	0.00
$\beta^*_{Hispanic}$	-0.87	0.15	-5.87	0.00
R^2		0.4178		
Adjusted R^2*		0.4154		
AIC*		47140.293		
BIC*		47247.021		

Note. *after accounting for weights and the design effect of 4

For MMM with unequal intercepts (MMMU, Table 4.B.6), the additive constant for the Hispanic ($\beta^*_{Hispanic}$) and the Black/African American ($\beta^*_{Black/African\ American}$) groups were no longer significant after accounting for varying intercepts between groups. On the other hand, the additive constant for the Asian group (β^*_{Asian}) became statistically significant after accounting for differences in intercepts. The β^*_{Asian} of -2.06 indicates that a one-point increase in total predictor score decreases the expected achievement by 2.06 points for the Asian group compared to the White group. When allowing differences in intercepts between groups, the Asian group, on average, had higher predicted mathematics achievement scores by 11.79 points than the White group. The

Black/African American group had lower predicted mean mathematics achievement score 9.81 points lower than the White group.

Table 4.B.6

Results for MMMU predicting mathematics achievement

	SE*	t*	p*	SE*
(Intercept)	58.23	1.13	51.75	0.00
Basic Math	-3.30	0.60	-5.47	0.00
Other Math	-1.89	0.51	-3.75	0.00
Pre-Algebra	-2.54	0.74	-3.41	0.00
Algebra I	-0.95	0.47	-2.04	0.04
Geometry	1.06	0.59	1.79	0.07
Algebra II	3.87	0.51	7.56	0.00
Trigonometry	7.01	0.66	10.61	0.00
Other Advanced	1.78	0.45	3.91	0.00
Probability and Statistics	4.94	1.00	4.94	0.00
Other AP/IB Math	8.07	1.03	7.87	0.00
Pre-Calculus	11.18	0.59	19.02	0.00
Calculus	10.85	1.05	10.32	0.00
AP/IB Calculus	13.55	0.72	18.92	0.00
Asian†	11.79	4.31	2.73	0.01
Black/African American†	-9.81	2.27	-4.32	0.00
Hispanic†	-2.69	1.79	-1.50	0.13
β^*_{Asian}	-2.06	1.00	-2.05	0.04
$\beta^*_{Black/African\ American}$	0.60	0.57	1.06	0.29
$\beta^*_{Hispanic}$	-0.21	0.47	-0.46	0.65
R^2		0.4221		
Adjusted R^2*		0.4193		
AIC*		47119.737		
BIC*		47245.299		

Note. *after accounting for weights and the design effect of 4; †the predictive mean difference (intercept) between this focal group and the reference (White) group

v) LMM

LMM allows regression weight vectors to differ by both additive and multiplicative constants, so this model is a combination of PMM and MMM. For LMM with a constant intercept (LMME. Table 4.B.7), additive constants were statistically significant for the Black/African American group (b = -1.61, p < .001) and the Hispanic group (b =-.70, p < .001). Neither additive nor multiplicative constants were significant for the Asian group of students.

Table 4.B.7

Results for LMME predicting mathematics achievement

	Estimate	SE*	t*	p*
(Intercept)	56.31	0.83	67.73	0.00
Basic Math	-2.95	0.59	-5.03	0.00
Other Math	-1.54	0.48	-3.17	0.00
Pre-Algebra	-2.19	0.73	-2.99	0.00
Algebra I	-0.77	0.44	-1.74	0.08
Geometry	1.61	0.56	2.85	0.00
Algebra II	4.50	0.48	9.43	0.00
Trigonometry	7.52	0.64	11.66	0.00
Other Advanced	2.27	0.41	5.47	0.00
Probability and Statistics	5.29	0.99	5.33	0.00
Other AP/IB Math	8.44	1.02	8.27	0.00
Pre-Calculus	11.82	0.58	20.35	0.00
Calculus	11.21	1.05	10.69	0.00
AP/IB Calculus	14.04	0.73	19.36	0.00
k_{Asian}	-0.09	0.10	-0.89	0.37
$k_{Black/African\ American}$	-0.06	0.08	-0.77	0.44
$k_{Hispanic}$	-0.07	0.06	-1.20	0.23
β^*_{Asian}	0.92	0.54	1.71	0.09
$\beta^*_{Black/African\ American}$	-1.61	0.23	-7.07	0.00

$\beta^*_{Hispanic}$	-0.70	0.20	-3.53	0.00
R^2		0.4181		
Adjusted R^2*		0.4153		
AIC*		47147.045		
BIC*		47272.607		

Note. *after accounting for weights and the design effect of 4

After allowing intercepts to differ by groups (LMMU), however, all the multiplicative constants (k_gs) and additive constants (β^*_gs) became statistically insignificant. The Asian group, on average, had higher predicted mathematics achievement scores 11.21 points higher than the White group, controlling for all other predictors. The Black/African American group, had, on average, mathematics achievement score 10.39 points lower than the White group.

Table 4.B.8

Results for LMMU predicting mathematics achievement

	Estimate	SE*	t*	p*
(Intercept)	58.56	1.14	51.42	0.00
Basic Math	-3.65	0.63	-5.80	0.00
Other Math	-2.19	0.53	-4.15	0.00
Pre-Algebra	-2.84	0.76	-3.74	0.00
Algebra I	-1.22	0.49	-2.50	0.01
Geometry	0.96	0.60	1.60	0.11
Algebra II	3.89	0.51	7.58	0.00
Trigonometry	7.07	0.66	10.69	0.00
Other Advanced	1.67	0.46	3.65	0.00
Probability and Statistics	4.92	1.00	4.92	0.00
Other AP/IB Math	8.19	1.03	7.97	0.00
Pre-Calculus	11.38	0.60	19.07	0.00
Calculus	10.87	1.05	10.35	0.00
AP/IB Calculus	13.76	0.74	18.71	0.00

Asian†	11.21	4.34	2.58	0.01
Black/African American†	-10.39	2.30	-4.52	0.00
Hispanic†	-3.40	1.84	-1.84	0.07
k_{Asian}	-0.08	0.10	-0.76	0.45
$k_{Black/African\ American}$	-0.11	0.08	-1.39	0.16
$k_{Hispanic}$	-0.09	0.06	-1.55	0.12
β^*_{Asian}	-1.62	1.15	-1.41	0.16
$\beta^*_{Black/African\ American}$	0.97	0.61	1.58	0.11
$\beta^*_{Hispanic}$	0.18	0.53	0.35	0.73
R^2		0.4227		
Adjusted R^2*		0.4195		
AIC*		47124.865		
BIC*		47269.261		

Note. *after accounting for weights and the design effect of 4; †the predictive mean difference (intercept) between this focal group and the reference (White) group

vi) FMM

FMM is the fullest model among all ten models suggested in this thesis, as it gives each group a distinct vector for regression coefficients without constraints. FMM with a constant intercept (FMME) is designed to model the situation in which everyone has the same starting point, such as the same average exam score in mathematics, but the relationship between each slope and the criterion may differ between subgroups. The slopes of all thirteen sequences of mathematics course-taking were statistically significant predictors of the post mathematics achievement of the White group without controlling for prior achievement. There were significant differences in slopes between the Black/African American group and the White group in Geometry (b = -5.08, p < .001) and Other Advanced (b = -2.48, p < .01) controlling for other variables. There was a significant difference in slopes between the Hispanic group and the White group in Geometry ($b = -2.56$, $p < .05$), controlling for other variables. In short, the relationship between Geometry and the criterion differs between the Black/African American or the Hispanic group and the White group: taking one more credit in Geometry is negatively associated with post-achievement for the Black/African American and the Hispanic group while it is positively associated with post-achievement for the White group.

Table 4.B.9
Results for FMME predicting mathematics achievement

	Estimate	SE*	t*	p*
(Intercept)	56.27	0.83	67.40	0.00
Basic Math	-3.10	0.73	-4.24	0.00
Other Math	-1.92	0.62	-3.12	0.00
Pre-Algebra	-2.49	0.94	-2.66	0.01
Algebra I	-1.23	0.53	-2.32	0.02
Geometry	2.45	0.70	3.49	0.00
Algebra II	4.49	0.61	7.42	0.00
Trigonometry	7.10	0.79	9.03	0.00
Other Advanced	2.49	0.52	4.84	0.00
Probability and Statistics	4.95	1.23	4.02	0.00
Other AP/IB Math	8.27	1.27	6.50	0.00
Pre-Calculus	11.78	0.70	16.88	0.00
Calculus	10.19	1.21	8.42	0.00
AP/IB Calculus	13.72	0.86	16.04	0.00
Asian:Basic Math	-1.72	4.45	-0.39	0.70
Asian:Other Math	-0.42	3.13	-0.14	0.89
Asian:Pre-Algebra	2.13	6.25	0.34	0.73
Asian:Algebra I	0.63	1.84	0.34	0.73
Asian:Geometry	4.04	2.73	1.48	0.14
Asian:Algebra II	-1.72	2.38	-0.72	0.47
Asian:Trigonometry	0.05	2.94	0.02	0.99
Asian:Other Advanced	1.63	1.96	0.83	0.41

Asian:Probability and Statistics	-6.24	6.07	-1.03	0.30
Asian:Other AP/IB Math	2.04	3.56	0.57	0.57
Asian:Pre-Calculus	-1.59	2.46	-0.65	0.52
Asian:Calculus	5.79	4.54	1.28	0.20
Asian:AP/IB Calculus	-0.63	2.00	-0.32	0.75
Black:Basic Math	0.00	1.50	0.00	1.00
Black:Other Math	0.08	1.05	0.08	0.94
Black:Pre-Algebra	0.59	1.74	0.34	0.74
Black:Algebra I	-1.59	1.21	-1.31	0.19
Black:Geometry	-5.08	1.42	-3.59	0.00
Black:Algebra II	0.31	1.35	0.23	0.82
Black:Trigonometry	-2.29	1.80	-1.27	0.20
Black:Other Advanced	-2.48	0.86	-2.87	0.00
Black:Probability and Statistics	1.17	2.84	0.41	0.68
Black:Other AP/IB Math	-3.11	3.90	-0.80	0.43
Black:Pre-Calculus	0.31	1.77	0.18	0.86
Black:Calculus	-3.88	4.23	-0.92	0.36
Black:AP/IB Calculus	-0.31	3.07	-0.10	0.92
Hispanic:Basic Math	-0.57	1.29	-0.44	0.66
Hispanic:Other Math	-0.13	1.22	-0.11	0.91
Hispanic:Pre-Algebra	-0.89	1.77	-0.50	0.61
Hispanic:Algebra I	1.08	0.84	1.29	0.20
Hispanic:Geometry	-2.56	1.15	-2.23	0.03
Hispanic:Algebra II	-1.63	1.06	-1.53	0.13
Hispanic:Trigonometry	1.94	1.70	1.14	0.26
Hispanic:Other Advanced	-1.64	1.04	-1.58	0.12
Hispanic:Probability and Statistics	-0.71	2.63	-0.27	0.79
Hispanic:Other AP/IB Math	-0.79	2.56	-0.31	0.76
Hispanic:Pre-Calculus	-1.95	1.45	-1.35	0.18
Hispanic:Calculus	4.30	3.19	1.35	0.18
Hispanic:AP/IB Calculus	-0.31	1.97	-0.16	0.88
R^2		0.4235		
Adjusted R^2*		0.4158		
AIC*		47209.763		
BIC*		47542.503		

Note. *after accounting for weights and the design effect of 4

When ethnic groups have different intercepts, as in the most complex model in this thesis, the full moderated model (FMMU) is used. FMMU does not give any constraints on ethnic groups and allows varying differences in slope and intercept for each focal group compared to the reference group. According to the results, the varying intercepts for the Asian and the Black/African American groups were significantly different from the intercept for the White group. The Asian group had predicted mean mathematics achievement score 11.50 points higher than the White group while the Black/African American group had predicted mean mathematics achievement score 9.06 points lower than the White group.

Table 4.B.10

Results for FMMU predicting mathematics achievement

	Estimate	SE*	t*	p*
(Intercept)	58.23	1.15	50.45	0.00
Basic Math	-3.60	0.76	-4.76	0.00
Other Math	-2.29	0.63	-3.62	0.00
Pre-Algebra	-2.99	0.96	-3.13	0.00
Algebra I	-1.74	0.57	-3.05	0.00
Geometry	1.78	0.75	2.38	0.02
Algebra II	4.06	0.63	6.45	0.00
Trigonometry	6.70	0.80	8.36	0.00
Other Advanced	2.01	0.55	3.66	0.00
Probability and Statistics	4.73	1.23	3.84	0.00
Other AP/IB Math	8.04	1.27	6.32	0.00
Pre-Calculus	11.44	0.71	16.14	0.00
Calculus	10.02	1.21	8.28	0.00
AP/IB Calculus	13.45	0.86	15.64	0.00
Asian†	11.50	4.77	2.41	0.02
Black/African American†	-9.06	2.41	-3.76	0.00

Hispanic†	-3.18	1.92	-1.66	0.10
Asian:Basic Math	-3.93	4.55	-0.86	0.39
Asian:Other Math	-1.17	3.14	-0.37	0.71
Asian:Pre-Algebra	-1.23	6.37	-0.19	0.85
Asian:Algebra I	-0.85	1.98	-0.43	0.67
Asian:Geometry	-0.13	3.20	-0.04	0.97
Asian:Algebra II	-4.79	2.67	-1.79	0.07
Asian:Trigonometry	-3.23	3.19	-1.01	0.31
Asian:Other Advanced	-1.12	2.26	-0.50	0.62
Asian:Probability and Statistics	-6.04	6.05	-1.00	0.32
Asian:Other AP/IB Math	0.75	3.60	0.21	0.84
Asian:Pre-Calculus	-3.85	2.61	-1.47	0.14
Asian:Calculus	3.36	4.61	0.73	0.47
Asian:AP/IB Calculus	-2.26	2.11	-1.07	0.28
Black:Basic Math	1.82	1.57	1.16	0.25
Black:Other Math	1.65	1.13	1.47	0.14
Black:Pre-Algebra	2.57	1.81	1.42	0.16
Black:Algebra I	1.42	1.45	0.98	0.33
Black:Geometry	-2.57	1.56	-1.64	0.10
Black:Algebra II	1.96	1.41	1.39	0.17
Black:Trigonometry	-0.24	1.88	-0.13	0.90
Black:Other Advanced	-0.34	1.03	-0.33	0.74
Black:Probability and Statistics	2.51	2.85	0.88	0.38
Black:Other AP/IB Math	-1.45	3.91	-0.37	0.71
Black:Pre-Calculus	1.73	1.81	0.96	0.34
Black:Calculus	-3.34	4.23	-0.79	0.43
Black:AP/IB Calculus	0.70	3.08	0.23	0.82
Hispanic:Basic Math	0.20	1.36	0.14	0.89
Hispanic:Other Math	0.42	1.26	0.34	0.74
Hispanic:Pre-Algebra	-0.03	1.87	-0.01	0.99
Hispanic:Algebra I	1.91	0.98	1.95	0.05
Hispanic:Geometry	-1.43	1.34	-1.06	0.29
Hispanic:Algebra II	-0.97	1.12	-0.87	0.39
Hispanic:Trigonometry	2.63	1.75	1.50	0.13
Hispanic:Other Advanced	-0.91	1.13	-0.81	0.42
Hispanic:Probability and Statistics	-0.33	2.63	-0.13	0.90
Hispanic:Other AP/IB Math	-0.42	2.56	-0.16	0.87
Hispanic:Pre-Calculus	-1.41	1.48	-0.95	0.34

Hispanic:Calculus	4.53	3.18	1.42	0.15
Hispanic:AP/IB Calculus	0.07	1.99	0.04	0.97
R^2		0.4269		
Adjusted R^2		0.4188		
AIC		47195.567		
BIC		47547.141		

Note. *after accounting for weights and the design effect of 4; †the predictive mean difference (intercept) between this focal group and the reference (White) group

vii) Model Comparisons

Model fit values are presented in table 4.B.11. In addition, all possible F-tests were performed to compare nested models using sum of squared errors and corresponding degrees of freedom for each model. The corresponding values and results are presented in table 4.B.12. According to the results of all the F-tests, LMMU was better than the other models. The adjusted R^2 favored LMMU while AIC and BIC favored the ANCOVA model.

PMME is significantly better than the base model ($F_{(3,3919)}$: 22.022, $p < .001$), which suggests that adding proportional slopes for each ethnic group helps to explain more variance in the mathematics achievement score. Comparing model fit values, PMME (B3) also has a higher adjusted R^2, and lower AIC and BIC values than the base model (B1). PMMU is significantly better than both ANCOVA ($F_{(3,3916)}$: 1.626, $p < .001$) and PMME ($F_{(3,3916)}$: 28.516, $p < .001$). Comparing model fit values, PMMU (B4) has a higher adjusted R^2 than both ANCOVA and PMME. The AIC and BIC values, however, favored ANCOVA over PMMU.

MMME is statistically better than the base model ($F_{(3,3919)}$: 41.333, p < .001), which suggests that adding additive slopes for each ethnic group helps to explain more variance in mathematics achievement. Comparing model fit values, MMME (B5) has a higher adjusted R^2, and lower AIC and BIC values than the base model (B1). MMMU is statistically better than both ANCOVA ($F_{(3,3916)}$: 2.064, p < .001) and MMME ($F_{(3,3916)}$: 9.844, p < .001). Comparing model fit values, MMMU (B6) has a higher adjusted R^2 than ANCOVA and MMME. AIC and BIC values, however, favored ANCOVA over MMMU. One can compare PMM and MMM using any model fit value since they have the same degrees of freedom (precluding a significance test). For example, MMME has a higher adjusted R-squared value, which means MMME is the better than PMME considering the same degrees of freedom in both models. Similarly, MMMU has a lower AIC value than PMMU.

LMME is significantly better than the base model ($F_{(6,3916)}$: 21.038, p < .001) and PMME ($F_{(3,3916)}$: 19.738, p < .001), which suggests that adding (multiplicative and) additive slopes for each ethnic group helps to explain more variance in mathematics achievement. LMMU is significantly better than either ANCOVA ($F_{(6,3913)}$: 1.552, p < .001), PMMU ($F_{(3,3913)}$: 1.727, p < .001), MMMU ($F_{(3,3913)}$:1.289, p < .001) or LMME

($F_{(3,3913)}$: 10.379, p < .001). Comparing model fit values, LMMU (B8) has a higher adjusted R^2 than the ANCOVA, PMMU, MMMU, and LMME. The AIC and BIC values, however, favored ANCOVA over LMMU.

FMME is significantly better than either PMME ($F_{(36,3883)}$: 2.654, p < .001), MMME ($F_{(36,3883)}$: 1.070, p < .05), or LMME ($F_{(33,3883)}$: 1.099, p < .01), which suggests that varying slopes for each predictor and ethnic group helps to explain more variance in mathematics achievement. FMMU is significantly better than FMME ($F_{(3,3880)}$: 7.659, p < .001). However, ANCOVA ($F_{(39,3880)}$: .984, p = .693), MMMU ($F_{(36,3880)}$: .894, p = 1.000), and LMMU ($F_{(33,3880)}$: 0.858, p = 1.000) are more parsimonious than FMMU. Since LMMU was significantly better than both ANCOVA ($F_{(6,3913)}$: 1.677, p < .001) and MMMU ($F_{(3,3913)}$: 1.289, p < .001), LMMU is the best model when predicting mathematics achievement with thirteen sequences of course-taking. LMMU (B8) also has the highest adjusted R^2 among all ten models. AIC and BIC values, however, favored ANCOVA over LMMU.

Table 4.B.11
Model fit values for each model

Model	R^2	**Adjusted** R^2	**AIC**	**BIC**
BASE	0.3994	0.3974	47253.889	47341.783
ANCOVA	0.4212	0.4189	**47116.942**	**47223.670**
PMME	0.4093	0.4069	47197.101	47303.829
PMMU	0.4219	0.4191	47121.056	47246.618
MMME	0.4178	0.4154	47140.293	47247.021

MMMU	0.4221	0.4193	47119.737	47245.299
LMME	0.4181	0.4153	47147.045	47272.607
LMMU	0.4227	**0.4195**	47124.865	47269.261
FMME	0.4235	0.4158	47209.763	47542.503
FMMU	**0.4269**	0.4188	47195.567	47547.141

Table 4.B.12

Results for F-tests of Model comparison

Models	Degrees of Freedom	F	p-value
Base vs. **ANCOVA**	(3, 3919)	49.352	<.001
Base vs. **PMME**	(3, 3919)	22.022	<.001
Base vs. **MMME**	(3, 3919)	41.333	<.001
Base vs. **LMME**	(6, 3916)	21.038	<.001
Base vs. **FMME**	(39, 3883)	4.169	<.001
ANCOVA vs. **PMMU**	(3, 3916)	1.626	<.001
ANCOVA vs. **MMMU**	(3, 3916)	2.064	<.001
ANCOVA vs. **LMMU**	(6, 3913)	1.677	<.001
ANCOVA vs. FMMU	(39, 3880)	0.984	0.693
PMME vs. **PMMU**	(3, 3916)	28.516	<.001
PMME vs. **LMME**	(3, 3916)	19.738	<.001
PMME vs. **FMME**	(36, 3883)	2.564	<.001
PMMU vs. **LMMU**	(3, 3913)	1.727	<.001
PMMU vs. FMMU	(36, 3880)	0.931	0.987
MMME vs. **MMMU**	(3, 3916)	9.844	<.001
MMME vs. LMME	(3, 3916)	0.750	1.000
MMME vs. **FMMU**	(36, 3883)	1.070	0.017
MMMU vs. **LMMU**	(3, 3913)	1.289	<.001
MMMU vs. FMMU	(36, 3880)	0.894	1.000
LMME vs. **LMMU**	(3, 3913)	10.379	<.001
LMME vs. **FMME**	(33, 3883)	1.099	0.002
LMMU vs. FMMU	(33, 3880)	0.858	1.000
FMME vs. **FMMU**	(3, 3880)	7.659	<.001

Note: the bolded model is statistically better or more parsimonious model, effective *N*: 3,936 (after accounting for weights and the design effect of 4)

C. Ethnic differences in the relationship between *Level* and *Pattern* of mathematics course-taking and mathematics achievement

Section C is predicting mathematics achievement using *Level* and *Pattern* as independent variables before adding prior achievement as a covariate. It is easier to understand and apply *Level* and *Pattern* when predictors are in common units. Here, Carnegie units completed in thirteen mathematics course sequences are the common units in predictors, and the *Level* is the mean of the Carnegie units completed and *Pattern* represents the content of course-taking sequences.

i) Base model

The base multiple regression model (Table 4.C.1) is the simplest model, and it assumes that all regression weights for all focal groups are equal to those of the reference group, "one size fits all." For this model, we estimated one optimal intercept and one optimal slope coefficient per predictor. The results suggest that *Level* and *Pattern* were positively related to mathematics achievement. The R^2 for Table 4.C.1 is the same as for Table 4.B.1 demonstrating the equivalence to the 2-predictor model containing *Level* and *Pattern* with the 13-predictor model containing the course categories as was proved in Davison and Davenport (2002).

Table 4.C.1

Results for the base model predicting mathematics achievement

	Estimate	SE*	t*	p*
(Intercept)	56.75	0.81	70.12	0.00
Level	54.33	2.58	21.04	0.00
Pattern	13.00	0.29	45.05	0.00
R^2		0.3994		
Adjusted R^2*		0.3990		
AIC*		47220.865		
BIC*		47239.700		

Note. *after accounting for weights and the design effect of 4

ii) ANCOVA

The ANCOVA model assumes the same regression slopes for all groups but allows for intercepts to vary by group. The results of ANCOVA show that all predictors as well as differences in intercepts for each group were statistically significant (See table 4.C.2). This model assumes that the slopes of predictors are the same regardless of group membership, but the intercepts differ between ethnic groups. The Asian group had the highest mean, on average, which was 3.26 points higher than the mean of the White group. Meanwhile, the Black/African American group and the Hispanic group, on average, had lower mathematics achievement scores (7.49 and 3.46 points, respectively) than the White group.

Table 4.C.2

Results for the ANCOVA model predicting mathematics achievement

	Estimate	SE*	t*	p*
(Intercept)	58.43	0.82	71.01	0.00
Level	53.25	2.55	20.90	0.00
Pattern	12.26	0.29	42.11	0.00
Asian†	3.26	1.21	2.69	0.01
Black/African American†	-7.49	0.69	-10.93	0.00
Hispanic†	-3.46	0.56	-6.14	0.00
R^2		0.4211		
Adjusted R^2*		0.4204		
AIC*		47084.459		
BIC*		47122.128		

Note. *after accounting for weights and the design effect of 4; †the predictive mean difference (intercept) between this focal group and the reference (White) group

iii) PMM

PMM constrains the difference in slopes between the focal group and the reference group to be a multiplicative constant of the slopes of the reference group. PMM with a constant intercept (PMME, Table 4.C.3) constrains differences in slopes between a focal group and the reference group without allowing intercepts to vary. The vector of regression weights for the White group is multiplied by -0.40 to get the difference in vectors of regression weights (*Level* and *Pattern*) between the Black/African American group and the White group. The vector of regression weights for the White group is multiplied by -0.19 to get the difference in vectors of regression weights (*Level* and *Pattern*) between the Hispanic group and the White group.

When the multiplicative constant, k_g, is between -1 and 0, the regression coefficients for the focal group are (1+k_g) times the regression coefficients for the reference group, where (1+k_g) is greater than 0 and less than 1. Therefore, the Black/African American group and the Hispanic group have smaller increment in the expected criterion than those of the White group.

Table 4.C.3

Results for PMME predicting mathematics achievement

	Estimate	SE*	t*	p*
(Intercept)	56.67	0.80	70.53	0.00
Level	57.70	2.63	21.91	0.00
Pattern	13.60	0.32	42.75	0.00
k_{Asian}	0.04	0.05	0.69	0.49
$k_{Black/African\ American}$	-0.40	0.06	-6.54	0.00
$k_{Hispanic}$	-0.19	0.04	-4.46	0.00
R^2		0.4083		
Adjusted R^2*		0.4075		
AIC*		47170.819		
BIC*		47208.488		

Note. *after accounting for weights and the design effect of 4

All multiplicative constants, k_gs, in PMM with unequal intercepts (PMMU) were not statistically significant after controlling for differences in intercepts (Table 4.C.4). The multiplicative constants for the Black/African American group and the Hispanic group in PMME were no longer statistically significant after controlling for differences in intercepts. All predictors for group differences in intercepts were statistically significant, and therefore all focal groups had different intercepts compared to the White group. The Asian group of students, on average, had higher predicted mean mathematics achievement scores by 5.73 points than the White group. The Black/African American group and the Hispanic group had lower predicted mean mathematics score (7.06 and 2.88 points lower, respectively) than the White group.

Table 4.C.4

Results for PMMU predicting mathematics achievement

	Estimate	SE*	t*	p*
(Intercept)	58.15	0.84	69.43	0.00
Level	54.73	2.67	20.50	0.00
Pattern	12.61	0.34	36.59	0.00
Asian†	5.73	1.97	2.90	0.00
Black/African American†	-7.06	0.83	-8.50	0.00
Hispanic†	-2.88	0.68	-4.23	0.00
k_{Asian}	-0.14	0.09	-1.68	0.09
$k_{Black/African\ American}$	-0.05	0.07	-0.72	0.47
$k_{Hispanic}$	-0.07	0.05	-1.41	0.16
R^2		0.4218		
Adjusted R^2*		0.4206		
AIC*		47089.099		
BIC*		47145.602		

Note. *after accounting for weights and the design effect of 4; †the predictive mean difference (intercept) between this focal group and the reference (White) group

iv) MMM

MMM constrains the difference in slopes between the focal group and the reference group to be a constant. When all regression coefficients between a focal group

g and the reference group differ by an additive constant, β_g^*, the vectors of regression coefficients between those two groups differ only in their means. MMM with a constant intercept (MMME, Table 4.C.5), the positive value for $\beta^*_{Black/African\ American}$ indicates that a one-point increase in total predictor score for the Black/African American group leads to a 3.69-point increase in the expected mathematics achievement score compared to the expected mathematics achievement score for the White group.

Table 4.C.5

Results for MMME predicting mathematics achievement

	Estimate	SE*	t*	p*
(Intercept)	56.53	0.81	69.82	0.00
Level	54.98	2.59	21.24	0.00
Pattern	12.37	0.37	33.58	0.00
β^*_{Asian}	0.73	1.12	0.65	0.52
$\beta^*_{Black/African\ American}$	3.69	0.85	4.32	0.00
$\beta^*_{Hispanic}$	0.52	0.69	0.75	0.45
R^2		0.4022		
Adjusted R^2*		0.4014		
AIC*		47211.111		
BIC*		47248.780		

Note. *after accounting for weights and the design effect of 4

In contrast, MMM with unequal intercepts (MMMU, Table 4.C.6) showed that the additive constant for the Black/African American became insignificant after accounting for group differences in intercepts. When checking the differences in intercepts between groups, the Asian group, on average, had higher predicted mathematics achievement scores by 3.73 points than the White group. The Black/African American group and the Hispanic group had lower predicted mean mathematics achievement score (7.99 and 3.75 points lower, respectively) lower than the White group, on average.

Table 4.C.6

Results for MMMU predicting mathematics achievement

	Estimate	SE*	t*	p*
(Intercept)	58.63	0.83	70.61	0.00
Level	53.21	2.55	20.83	0.00
Pattern	12.70	0.36	34.85	0.00
Asian†	3.73	1.32	2.83	0.00
Black/African American†	-7.99	0.81	-9.84	0.00
Hispanic†	-3.75	0.61	-6.13	0.00
β^*_{Asian}	-1.53	1.20	-1.27	0.20
$\beta^*_{Black/African\ American}$	-1.35	0.99	-1.36	0.17
$\beta^*_{Hispanic}$	-1.03	0.74	-1.40	0.16
R^2	0.4217			
Adjusted R^2*	0.4206			
AIC*	47089.272			
BIC*	47145.775			

Note. *after accounting for weights and the design effect of 4; †the predictive mean difference (intercept) between this focal group and the reference (White) group

v) LMM

LMM (allows regression weight vectors to differ by both additive and multiplicative constants, so this model is a combination of PMM and MMM. For LMM with a constant intercept (LMME, Table 4.C.7), both a multiplicative constant (k_g) and an additive constant (β_g^*) were statistically significant for the Black/African American group and the Hispanic group. The difference in vectors of regression coefficients differs by -.57*β_{v0}+6.64 for the Black/African American group while the corresponding difference is -.29*β_{v0}+3.02 for the Hispanic group, where β_{v0} is a vector of regression coefficients for the White group of students. For the Asian group, only the multiplicative constant was statistically significant (b = .205, p < .05).

Table 4.C.7

Results for LMME predicting mathematics achievement

	Estimate	SE*	t*	p*
(Intercept)	56.40	0.80	70.50	0.00
Level	59.95	2.64	22.72	0.00
Pattern	12.49	0.36	34.31	0.00
k_{Asian}	0.20	0.10	2.04	0.04
$k_{Black/African\ American}$	-0.57	0.06	-8.75	0.00
$k_{Hispanic}$	-0.29	0.05	-5.85	0.00
β^*_{Asian}	-3.79	2.13	-1.78	0.07
$\beta^*_{Black/African\ American}$	6.64	0.90	7.36	0.00
$\beta^*_{Hispanic}$	3.02	0.79	3.82	0.00
R^2		0.4179		
Adjusted R^2*		0.4168		
AIC*		47115.149		
BIC*		47171.652		

Note. *after accounting for weights and the design effect of 4

Furthermore, when intercepts were allowed to differ by group (LMMU, Table 4.C.8), both multiplicative and additive constants were insignificant for all focal groups of students. The Asian group had, on average, higher predicted mathematics achievement score (11.17 points higher) than the White group, controlling for all other predictors. The Black/African American group tended to have lower predicted scores in mathematics achievement by 10.32 points than the White group.

Table 4.C.8
Results for LMMU predicting mathematics achievement

	Estimate	SE*	t*	p*
(Intercept)	58.61	1.11	52.63	0.00
Level	53.28	3.52	15.12	0.00
Pattern	12.70	0.37	34.32	0.00
Asian†	11.17	4.34	2.58	0.01
Black/African American†	-10.32	2.29	-4.50	0.00
Hispanic†	-3.34	1.85	-1.81	0.07
k_{Asian}	-0.59	0.33	-1.79	0.07
$k_{Black/African\ American}$	0.20	0.18	1.10	0.27
$k_{Hispanic}$	-0.04	0.15	-0.24	0.81
β^*_{Asian}	6.56	4.67	1.41	0.16
$\beta^*_{Black/African\ American}$	-3.76	2.48	-1.52	0.13
$\beta^*_{Hispanic}$	-0.56	2.13	-0.26	0.79
R^2	0.4225			
Adjusted R^2*	0.4209			
AIC*	47093.142			
BIC*	47168.479			

Note. *after accounting for weights and the design effect of 4; †the predictive mean difference (intercept) between this focal group and the reference (White) group

vi) FMM

FMM with a constant intercept (FMME, Table 4.C.9) is designed to model the situation in which everyone has the same starting point, such as same average score in mathematics, but there are different slopes for each predictor between subgroups. Both the Black/African American group and the Hispanic group have statistically different regression coefficients for *Level* compared to the White group. However, there was not enough evidence for differences in *Pattern* compared to the White group.

Table 4.C.9

Results for FMME predicting mathematics achievement

	Estimate	SE*	t*	p*
(Intercept)	56.40	0.80	70.50	0.00
Level	59.95	2.64	22.72	0.00
Pattern	12.49	0.36	34.31	0.00
Asian:Level	7.11	3.72	1.91	0.06
Asian:Pattern	-1.19	1.24	-0.97	0.33
Black:Level	-23.60	3.26	-7.25	0.00
Black:Pattern	-0.57	0.99	-0.58	0.56
Hispanic:Level	-12.54	2.37	-5.29	0.00
Hispanic:Pattern	-0.69	0.73	-0.95	0.34
R^2		0.4179		
Adjusted R^2*		0.4168		
AIC*		47115.149		
BIC*		47171.652		

Note. *after accounting for weights and the design effect of 4

The full moderated model (FMMU, Table 4.C.10) is the most complex model among the ten models. According to the results, for all focal groups, neither Level nor Pattern were significantly different from that for the White group. The Asian group had predicted mean mathematics score 11.17 points higher than the White group, while the

Black/African American group had predicted mean mathematics scores 10.32 points lower than the White group after controlling for the effect of *Level* and *Pattern*.

Table 4.C.10

Results for FMMU predicting mathematics achievement

	Estimate	SE*	t*	p*
(Intercept)	58.61	1.11	52.63	0.00
Level	53.28	3.52	15.12	0.00
Pattern	12.70	0.37	34.32	0.00
Asian†	11.17	4.34	2.58	0.01
Black/African American†	-10.32	2.29	-4.50	0.00
Hispanic†	-3.34	1.85	-1.81	0.07
Asian:Level	-25.03	13.15	-1.90	0.06
Asian:Pattern	-0.97	1.24	-0.78	0.43
Black:Level	6.92	7.52	0.92	0.36
Black:Pattern	-1.21	1.00	-1.22	0.22
Hispanic:Level	-2.47	6.05	-0.41	0.68
Hispanic:Pattern	-1.02	0.75	-1.36	0.17
R^2	0.4225			
Adjusted R^2*	0.4209			
AIC*	47093.142			
BIC*	47168.479			

Note. *after accounting for weights and the design effect of 4; †the predictive mean difference (intercept) between this focal group and the reference (White) group

vii) Model Comparisons

The ANCOVA model is significantly better than the base model ($F_{(3,3930)}$: 49.301, p < .001). There are group mean differences to consider, and one size does not fit all. The ANCOVA model (C2) has a higher adjusted R^2, and lower AIC and BIC values than the base model (C1).

PMME is significantly better than the base model ($F_{(3,3930)}$: 19.801, p < .001), which suggests that adding proportional slopes for each ethnic group helps to explain more variance in mathematics achievement. PMMU is significantly better than both PMME ($F_{(3,3927)}$: 30.523, p < .001) and ANCOVA ($F_{(3,3927)}$: 1.453, p < .001).

MMME is statistically better than the base model ($F_{(3,3930)}$: 6.258, p < .001), which suggests that adding additive slopes for each ethnic group helps to explain more variance in mathematics achievement. MMMU is statistically better than both ANCOVA ($F_{(3,3927)}$: 1.395, p < .001) and MMME ($F_{(3,3927)}$: 44.246, p < .001). PMME is better than MMME since it has a higher adjusted R^2 value. PMMU is better than MMMU since it has a lower BIC value.

LMME is significantly better than the base model ($F_{(6,3927)}$: 20.900, p < .001), which suggests that adding both proportional and additive slopes for each ethnic group helps explain more variance in mathematics achievement. LMME is significantly better than both PMME ($F_{(3,3927)}$: 21.687, p < .001) and MMME ($F_{(3,3927)}$: 35.379, p < .001). LMMU is significantly better than either ANCOVA ($F_{(3,3924)}$: 1.552, p < .001), PMMU ($F_{(3,3924)}$: 1.651, p < .011) or LMME ($F_{(3,3924)}$: 10.348, p < .001).

FMME was better than either the base model ($F_{(6,3927)}$:20.900, p < .001) or PMME ($F_{(3,3927)}$: 21.687, p < .001) or MMME ($F_{(3,3927)}$: 35.379, p < .001). FMMU was better than either ANCOVA ($F_{(6,3924)}$: 1.552, p < .001) or PMMU ($F_{(3,3924)}$: 1.651,

p < .011) or FMME ($F_{(3,3924)}$: 10.348, p < .001). We cannot compare LMMU and FMMU using the F-test, since the degrees of freedom are the same for both models. In addition, all model fit values were the same for LMMU and FMMU.

According to the results of all the F-tests for comparing nested models, LMMU was better than the other models. The adjusted R^2 favored LMMU (sharing the same model fit values with FMMU), while AIC and BIC favored the ANCOVA model.

Table 4.C.11

Model fit values for each model

Model	R^2	**Adjusted** R^2	**AIC**	**BIC**
BASE	0.3994	0.3990	47220.865	47239.700
ANCOVA	0.4211	0.4204	**47084.459**	**47122.128**
PMME	0.4083	0.4075	47170.819	47208.488
PMMU	0.4218	0.4206	47089.099	47145.602
MMME	0.4022	0.4014	47211.111	47248.780
MMMU	0.4217	0.4206	47089.272	47145.775
LMME	0.4179	0.4168	47115.149	47171.652
LMMU	**0.4225**	**0.4209**	47093.142	47168.479
FMME	0.4179	0.4168	47115.149	47171.652
FMMU	**0.4225**	**0.4209**	47093.142	47168.479

Table 4.C.12

Results for F-tests of Model comparison

Models	Degrees of Freedom	F	p-value
Base vs. **ANCOVA**	(3, 3930)	49.301	<.001
Base vs. **PMME**	(3, 3930)	19.801	<.001
Base vs. **MMME**	(3, 3930)	6.258	<.001
Base vs. **LMME**	(6, 3927)	20.900	<.001
Base vs. **FMME**	(6, 3927)	20.900	<.001

ANCOVA vs. **PMMU**	(3, 3927)	1.453	<.001
ANCOVA vs. **MMMU**	(3, 3927)	1.395	<.001
ANCOVA vs. **LMMU**	(6, 3924)	1.552	<.001
ANCOVA vs. **FMMU**	(6, 3924)	1.552	<.001
PMME vs. **PMMU**	(3, 3927)	30.523	<.001
PMME vs. **LMME**	(3, 3927)	21.687	<.001
PMME vs. **FMME**	(3, 3927)	21.687	<.001
PMMU vs. **LMMU**	(3, 3924)	1.651	<.001
PMMU vs. **FMMU**	(3, 3924)	1.651	<.001
MMME vs. **MMMU**	(3, 3927)	44.246	<.001
MMME vs. **LMME**	(3, 3927)	35.379	<.001
MMME vs. **FMME**	(3, 3924)	35.379	<.001
MMMU vs. **LMMU**	(3, 3924)	1.709	<.001
MMMU vs. **FMMU**	(3, 3924)	1.709	<.001
LMME vs. **LMMU**	(3, 3924)	10.348	<.001
LMME vs. FMME	-	-	-
LMMU vs. FMMU	-	-	-
FMME vs. **FMMU**	(3, 3924)	10.348	<.001

Note: the bolded model is statistically better or more parsimonious model, effective *N*: 3,936 (after accounting for weights and the design effect of 4)

D. Possible covariates: prior achievement and socio-economic status (SES)

i) Prior achievement

Prior mathematics achievement is the mathematics score at the time of the base-year assessment in mathematics (X2X1TXMSCR). It is the re-estimated score that students would have answered correctly had they responded to all 118 items in the 9th grade. Adding prior mathematics achievement explains 20.3% more variance in post mathematics achievement after controlling for *Level* and *Pattern*. Once controlling for prior achievement, the effects of *Level* and *Pattern* on the mathematics achievement persist, but at a much lower magnitude.

Table 4.D.1

Regression analyses predicting mathematics achievement using Level and Pattern (Model 1), and adding Prior Achievement (Model 2)

Independent Variables	Model 1	Model 2
(Intercept)	56.752*** (0.405)	23.012*** (0.500)
Level	54.328*** (1.291)	27.836*** (1.092)
Pattern	13.000*** (0.144)	5.851*** (0.142)
Prior Achievement		0.682*** (0.008)
R^2	0.399	0.602
R^2 change		0.203

Note: Numbers are unstandardized regression weights (standard errors) except for R^2 and R^2 change; *** p < .001 after accounting for design effect of 4.

ii) Socio-economic status (SES)

Two slightly different indices for socio-economic status (SES) were developed for HSLS:09 using multiple imputation (MI). The first index (X1SES) included responses from the parent questionnaire for all parent/guardian types including parent education, parent occupation, and family income as well as student home-life (contextual) analysis weight and HSLS:09 sample design. A second SES index (X1SES_U) was a variant of X1SES, and it accounts for differences in the target population by school urbanicity (Ingels et al., 2011). Both variables were included separately to assess the need for SES as a covariate in this thesis even after controlling for prior achievement. While X1SES_U is recommended for use in a bivariate or multivariate analysis, X1SES is better for use in NCES secondary longitudinal studies (Ingels et al., 2011), since these studies did not account for locale (urbanicity) when calculating the SES variable in the past. When multivariate analysis controls for locale, it is okay to use X1SES since the locale variable has already been considered. Consequently, SES with locale (X1SES_U) is a more reliable measure of SES, since it covers information about urbanicity as well as SES.

First, the multiple regression was performed to predict mathematics achievement by prior achievement, *Level*, and *Pattern*. This was followed by multiple regression with ‘SES’ added to the prior model. This entails regressing mathematics achievement scores on prior achievement, *Level*, *Pattern*, and SES, which was performed to assess the incremental effect of SES. The incremental effect on R^2 when adding X1SES to the prior model was .005 ($p < .001$), which indicates that SES explains 0.5% more variance in mathematics achievement. This is a small change, even though SES was a statistically significant predictor in the model. We should not consider this to be a practical difference although there is a statistical difference between the two regression models.

The same procedure with SES with locale instead of SES alone was repeated. I performed multiple regression to predict mathematics achievement by prior achievement, *Level*, and *Pattern*. This was followed by a fuller multiple regression to model mathematics achievement based on prior achievement, *Level*, *Pattern*, and SES with locale (urbanicity), which was performed to assess the incremental effect of SES with locale (X1SES_U). The incremental R^2 when adding X1SES_U to the prior model was .005 ($p < .001$), which indicates that SES explains 0.5% more variance in mathematics achievement. Considering the practical difference of 0.5% of the variance in mathematics achievement, SES does not seem to add explanatory power to the model.

Moreover, prior achievement will be controlled, which is also related to SES or SES with locale ($r = .419$ or $r = .413$, respectively), while the correlation between the SES or SES with locale and post achievement is 0.412 or 0.407, respectively. In addition, prior achievement is used as a surrogate measure for anything relative to mathematics achievement prior to the base year, as Hanushek (1972) recommended the use of prior achievement to capture past ability and learning environment.

Table 4.D.2

Regression analyses predicting mathematics achievement using prior achievement, Level and Pattern (Model 1), and adding SES (Model 2a) or SES with locale (Model 2b)

Independent Variables	Model 1	Model 2a)	Model 2b)
(Intercept)	23.012***	24.885***	24.839***
Prior Achievement	0.682***	0.656***	0.657***
Level	27.836***	26.074***	26.108***
Pattern	5.851***	5.524***	5.533***
2a) SES		1.881***	
2b) SES with locale (urbanicity)			1.853***
R^2	0.602	0.607	0.607
R^2 change		0.005	0.005

Note: all numbers are unstandardized regression weights, and *** p < .001 after accounting for design effect of 4.

E. Ethnic differences in the relationship between prior achievement, thirteen sequences of mathematics course-taking and mathematics achievement

In this section, mathematics achievement is predicted by prior achievement and credits for each of thirteen course sequences (X1: Basic Math, X2: Other Math, X3: Pre-Algebra, X4: Algebra I, X5: Geometry, X6: Algebra II, X7: Trigonometry, X8: Other advanced math, X9: Probability and statistics, X10: Other AP/IB math, X11: Precalculus, X12: Calculus, and X13: AP/IB Calculus), and moderated by ethnic group (White is the referent group, $Z_1 = 1$ for Asian, $Z_2 = 1$ for Black/African American, and $Z_3 = 1$ for Hispanic).

i) Base model

The base multiple regression model (Table 4.E.1) is the simplest model, and it assumes "one size fits all." For this model, we estimated one optimal intercept and one optimal slope coefficient per predictor. The results imply that, on average, credits in basic mathematics courses categorized as "Basic Math" or "Other Math" were negatively related to mathematics achievement, controlling for prior achievement and other course sequences. Taking one more credit in Pre-Algebra, Algebra I, and Geometry does not make a statistically significant difference in future achievement, controlling for prior achievement and other course sequences. However, higher level courses, including "Calculus" or "AP/AB Calculus," are positively related to mathematics achievement. If we rearrange predictors in order of increasing regression coefficient value, Other advanced math has the sixth lowest partial regression coefficient, followed by Probability and statistics, Algebra II, Other AP/IB math, Trigonometry, Calculus, Precalculus, and AP/IB Calculus. Controlling for other variables, taking one more credit in AP/IB Calculus is associated with 6.62 points increase in predicted mathematics achievement.

Table 4.E.1

Results for the base model predicting mathematics achievement

	Estimate	SE*	t*	p*
(Intercept)	22.66	1.02	22.18	0.00
Prior Achievement	0.69	0.02	44.94	0.00
Basic Math	-1.38	0.47	-2.94	0.00
Other Math	-0.98	0.38	-2.54	0.01
Pre-Algebra	-0.02	0.59	-0.03	0.98
Algebra I	-0.01	0.35	-0.02	0.99
Geometry	0.45	0.46	0.99	0.32
Algebra II	2.07	0.39	5.31	0.00
Trigonometry	3.86	0.53	7.26	0.00
Other Advanced	0.75	0.34	2.22	0.03
Probability and Statistics	1.80	0.82	2.19	0.03
Other AP/IB Math	3.10	0.84	3.68	0.00
Pre-Calculus	5.88	0.48	12.17	0.00
Calculus	5.47	0.87	6.26	0.00
AP/IB Calculus	6.62	0.59	11.22	0.00
R^2		0.6031		
Adjusted R^2*		0.6017		
AIC*		45625.334		
BIC*		45719.506		

Note. *after accounting for weights and the design effect of 4

Comparing this model (E1) with the model without prior achievement (B1), all absolute values for coefficients in the prior model (B1) are greater than this model (E1) with prior achievement. For example, taking one more credit in AP/AB Calculus is positively associated with a 14.13-point increase in mathematics achievement, while the same event is positively associated with only a 6.62-point increase after accounting for prior achievement. Note that prior achievement partially mediates the relationship of the other predictors.

Unsurprisingly, the order of regression coefficients was also different from the order of regression coefficients when accounting for prior achievement. These differences in the order of regression coefficients in both base models are related to the correlation between the predictor and a covariate (prior achievement) as well as correlation of predictors by themselves. For example, "Probability and Statistics (b = 5.02)" had a larger regression coefficient than "Algebra II (b = 4.00)," while the order of their regression coefficients was the opposite (b = 1.80 and 2.07, respectively) when considering the effect of prior achievement. Interestingly, the strength of overall correlation (without considering group membership) between "Probability and Statistics" and prior achievement is greater than the correlation between "Probability and Statistics" and post achievement (r =.087 and .078, respectively). The strength of overall correlation between "Algebra II" and prior achievement, however, is lower than the correlation between "Algebra II" and post achievement (r =.170 and .182, respectively).

In addition, “Other AP/AB Math (b = 8.23)” was more strongly related to the criterion than “Trigonometry (b = 7.10)” while the order of their regression coefficients was the opposite (b = 3.10 and 3.86, respectively) when considering the effect of prior achievement. Similarly, the strength of overall correlation between “Other AP/AB Math” and prior achievement is greater than the correlation between “Other AP/AB Math” and post achievement (r =.234 and .228, respectively). The strength of overall correlation between “Trigonometry” and prior achievement, however, is lower than the correlation between “Trigonometry” and post achievement (r =.164 and .186, respectively). As in the same way, “Other Math (b = -2.83)” had a larger absolute value of regression coefficient than that of “Pre-Algebra (b = -2.21),” while the order was the opposite when considering the effect of prior achievement (b = -.02 and -.98, respectively). Interestingly, the strength of the correlation between “Pre-Algebra” and prior achievement is greater than the correlation between “Pre-Algebra” and post achievement (r =-.206 and -.180, respectively). The strength of the correlation between “Other Math” and prior achievement, however, is lower than the correlation between “Other Math” and post achievement (r =-.147 and -.161, respectively). It is notable that prior achievement partially mediates regression coefficients of other predictors, so that the size and order of regression coefficients differs before and after accounting for prior achievement.

ii) ANCOVA

The ANCOVA model (Table 4.E.2) assumes the same regression slopes for all groups but allows for different intercepts. In doing so, this model tests whether the means differ between ethnic groups while controlling for the effect of other predictors . It is found that the means of Black/African American and Hispanic groups statistically differ from that of the reference group, the White group. The Black/African American and the Hispanic groups had lower predicted mean mathematics scores (2.89 and 1.30 points lower, respectively) than the White group of students. Also, having smaller absolute values of regression coefficients compared to the model without prior achievement (Table 4.B.2) is understandable since prior achievement accounts for some variance in mathematics achievement.

The order of regression coefficients for course-taking sequences are the same as for the base model: Basic Math, Other Math, Pre-Algebra, Algebra I, Geometry, Other Advanced, Probability and Statistics, Algebra II, Other AP/IB Math, Trigonometry, Calculus, Pre-Calculus, and AP/IB Calculus. As in the base model, Basic and Other Math were negatively related to mathematics achievement while higher mathematics course-taking was positively related to mathematics achievement given what is in the model.

Table 4.E.2

Results for the ANCOVA model predicting mathematics achievement

	Estimate	SE*	t*	p*
(Intercept)	24.16	1.06	22.76	0.00
Prior Achievement	0.67	0.02	43.00	0.00
Basic Math	-1.40	0.47	-2.97	0.00
Other Math	-0.89	0.38	-2.32	0.02
Pre-Algebra	0.02	0.59	0.04	0.97
Algebra I	0.06	0.35	0.17	0.86
Geometry	0.54	0.46	1.19	0.24
Algebra II	2.07	0.39	5.31	0.00
Trigonometry	3.91	0.53	7.34	0.00
Other Advanced	0.84	0.34	2.50	0.01
Probability and Statistics	1.84	0.82	2.24	0.02
Other AP/IB Math	3.09	0.84	3.68	0.00
Pre-Calculus	5.85	0.48	12.14	0.00
Calculus	5.34	0.87	6.12	0.00
AP/IB Calculus	6.49	0.59	10.93	0.00
Asian†	1.56	1.01	1.54	0.12
Black/African American†	-2.89	0.58	-5.01	0.00
Hispanic†	-1.30	0.47	-2.77	0.01
R^2		0.6063		
Adjusted R^2*		0.6046		
AIC*		45602.583		
BIC*		45715.589		

Note. *after accounting for weights and the design effect of 4; †the predictive mean difference (intercept) between this focal group and the reference (White) group

iii) PMM

PMM with a constant intercept (PMME, Table 4.E.3) constrains slope differences between groups without varying group intercepts. The vector of regression weights for the White group is multiplied by -.07 to get the difference in vectors of regression weights (prior achievement, Basic Math to AP/IB Calculus) between the Black/African American group and the White group. The vector of regression weights for the White group is multiplied by -.03 to get the difference in vectors of regression weights (prior achievement, Basic Math to AP/IB Calculus) between the Hispanic group and the White group.

When the multiplicative constant, k_g, is between -1 and 0, the regression coefficients for the focal group are (1+k_g) times the regression coefficients for the reference group, where (1+k_g) is greater than 0 and less than 1. Therefore, each focal group, here the Black/African American or the Hispanic, has a smaller increment in the expected Y than those of the White group, since the absolute values of the regression coefficients for these focal groups are smaller compared to those for the regression coefficients for the reference group.

Table 4.E.3

Results for PMME predicting mathematics achievement

	Estimate	SE*	t*	p*
(Intercept)	23.33	1.03	22.75	0.00
Prior Achievement	0.68	0.02	44.42	0.00
Basic Math	-1.41	0.47	-2.99	0.00
Other Math	-0.92	0.38	-2.40	0.02
Pre-Algebra	0.03	0.59	0.05	0.96
Algebra I	0.07	0.35	0.19	0.85
Geometry	0.57	0.46	1.25	0.21
Algebra II	2.14	0.39	5.50	0.00
Trigonometry	3.96	0.53	7.44	0.00

Other Advanced	0.88	0.34	2.61	0.01
Probability and Statistics	1.85	0.82	2.25	0.02
Other AP/IB Math	3.10	0.84	3.68	0.00
Pre-Calculus	5.93	0.48	12.31	0.00
Calculus	5.29	0.87	6.07	0.00
AP/IB Calculus	6.45	0.60	10.78	0.00
k_{Asian}	0.02	0.02	1.13	0.26
$k_{Black/African\ American}$	-0.07	0.01	-4.98	0.00
$k_{Hispanic}$	-0.03	0.01	-2.62	0.01
R^2		0.6062		
Adjusted R^2*		0.6045		
AIC*		45603.960		
BIC*		45716.966		

Note. *after accounting for weights and the design effect of 4

All multiplicative constants (k_gs) as well as differences in intercepts in PMM with unequal intercepts (PMMU) were not statistically significant after controlling for differences in intercepts (Table 4.E.4).

Table 4.E.4

Results for PMMU predicting mathematics achievement

	Estimate	SE*	t*	p*
(Intercept)	23.75	1.18	20.05	0.00
Prior Achievement	0.67	0.02	38.83	0.00
Basic Math	-1.39	0.47	-2.96	0.00
Other Math	-0.90	0.38	-2.35	0.02
Pre-Algebra	0.04	0.59	0.06	0.95
Algebra I	0.06	0.35	0.18	0.86
Geometry	0.55	0.46	1.19	0.23
Algebra II	2.09	0.39	5.31	0.00
Trigonometry	3.93	0.53	7.35	0.00
Other Advanced	0.86	0.34	2.53	0.01

Probability and Statistics	1.84	0.82	2.25	0.02
Other AP/IB Math	3.14	0.84	3.72	0.00
Pre-Calculus	5.88	0.48	12.15	0.00
Calculus	5.34	0.87	6.12	0.00
AP/IB Calculus	6.56	0.60	10.86	0.00
Asian†	4.40	3.56	1.24	0.22
Black/African American†	-1.43	1.81	-0.79	0.43
Hispanic†	-0.95	1.50	-0.63	0.53
k_{Asian}	-0.05	0.06	-0.84	0.40
$k_{Black/African\ American}$	-0.04	0.05	-0.84	0.40
$k_{Hispanic}$	-0.01	0.04	-0.22	0.82
R^2		0.6065		
Adjusted R^2*		0.6045		
AIC*		45610.289		
BIC*		45742.129		

Note. *after accounting for weights and the design effect of 4; †the predictive mean difference (intercept) between this focal group and the reference (White) group

iv) MMM

MMM constrains the difference in slopes between a focal group and the reference group to be a constant. When all regression coefficients between a focal group g and a reference group differ by an additive constant, β_g^*, the vectors of regression coefficients between those two groups differ only in their means. For MMM with a constant intercept (MMME, Table 4.E.5), the negative values of $\beta^*_{Black/African\ American}$ and $\beta^*_{Hispanic}$ indicate lower regression coefficients for the Black/African American (-.05) and the Hispanic (-.02) groups, in which a one-point increase in total predictor score leads to a 0.05-point and a 0.02-point decrease in the expected mathematics achievement scores compared to the expected mathematics achievement score for the White group. Taking the same total number of courses is not worth as much for the Hispanic and Black groups as for the White group. This could also be because of the types of courses students are taking.

Table 4.E.5

Results for MMME predicting mathematics achievement

	Estimate	SE*	t*	p*
(Intercept)	23.35	1.03	22.76	0.00
Prior Achievement	0.68	0.02	44.43	0.00
Basic Math	-1.36	0.47	-2.90	0.00
Other Math	-0.87	0.38	-2.27	0.02
Pre-Algebra	0.05	0.59	0.09	0.93
Algebra I	0.08	0.35	0.23	0.81
Geometry	0.56	0.46	1.22	0.22
Algebra II	2.11	0.39	5.42	0.00
Trigonometry	3.92	0.53	7.36	0.00
Other Advanced	0.87	0.34	2.59	0.01
Probability and Statistics	1.84	0.82	2.25	0.02
Other AP/IB Math	3.06	0.84	3.64	0.00
Pre-Calculus	5.86	0.48	12.15	0.00
Calculus	5.27	0.87	6.04	0.00
AP/IB Calculus	6.40	0.60	10.75	0.00
β^*_{Asian}	0.02	0.01	1.35	0.18
$\beta^*_{Black/African\ American}$	-0.05	0.01	-5.04	0.00
$\beta^*_{Hispanic}$	-0.02	0.01	-2.59	0.01
R^2		0.6063		
Adjusted R^2*		0.6046		
AIC*		45602.964		
BIC*		45715.970		

Note. *after accounting for weights and the design effect of 4

For MMM with unequal intercepts (MMMU, Table 4.E.6), the additive constants for the Black/African American group and the Hispanic group were no longer statistically significant after accounting for group differences in intercepts. On the other hand, the additive constant for the Asian group (β^*_{Asian}) became statistically significant after accounting for differences in intercepts compared to the White group. Here, the β^*_{Asian} of .30 indicates that a one-point increase in total predictor score increases the expected achievement by .30 points for the Asian group compared to the White group. Taking the same total number of courses is worth more for the Asian group than the White group when accounting for prior achievement (E6) while taking the same total number of courses is worth less for the Asian group (β^*_{Asian}= -2.06) compared to the White group without accounting for prior achievement (B6). In addition, all other significant differences in intercepts in MMMU became insignificant after accounting for prior achievement (B6 vs E6). This suggests that prior achievement explains some variance in group means, and it contains some information for the course-taking also.

All differences in intercepts between a focal group and the reference group were not statistically significant in MMMU. This is also confirmed with results from model comparison. When comparing MMME with MMMU, MMME is more parsimonious than MMMU ($F_{(3,3915)}$: 0.441, p = 1.000), which means allowing intercepts to vary for each focal group does not statistically improve the MMME model.

Table 4.E.6

Results for MMMU predicting mathematics achievement

	Estimate	SE*	t*	p*
(Intercept)	23.43	1.03	22.82	0.00
Prior Achievement	0.68	0.02	43.72	0.00
Basic Math	-1.38	0.48	-2.91	0.00
Other Math	-0.89	0.39	-2.25	0.02
Pre-Algebra	0.07	0.59	0.11	0.91
Algebra I	0.03	0.36	0.08	0.94
Geometry	0.54	0.46	1.17	0.24
Algebra II	2.16	0.40	5.41	0.00
Trigonometry	3.96	0.54	7.33	0.00
Other Advanced	0.87	0.34	2.58	0.01
Probability and Statistics	1.89	0.82	2.31	0.02
Other AP/IB Math	3.19	0.85	3.77	0.00
Pre-Calculus	5.96	0.51	11.74	0.00
Calculus	5.35	0.87	6.12	0.00
AP/IB Calculus	6.73	0.62	10.78	0.00
Asian†	-0.37	0.19	-1.91	0.06
Black/African American†	0.03	0.14	0.22	0.83
Hispanic†	-0.06	0.10	-0.61	0.54
β^*_{Asian}	0.30	0.15	2.03	0.04
$\beta^*_{Black/African\ American}$	-0.07	0.10	-0.74	0.46
$\beta^*_{Hispanic}$	0.02	0.07	0.33	0.74
R^2	0.6064			
Adjusted R^2*	0.6044			
AIC*	45610.649			
BIC*	45742.489			

Note. *after accounting for weights and the design effect of 4; †the predictive mean difference (intercept) between this focal group and the reference (White) group

v) LMM

LMM allows regression weight vectors to differ by both additive and multiplicative constants. This model is a combination of PMM and MMM. For LMM with a constant intercept (LMME, Table 4.E.7), an additive constant (β^*_{Asian}) was only statistically significant for the Asian group (b = 0.30, p < .05). However, neither additive nor multiplicative constants were significant for the Black/African American group and the Hispanic group.

Table 4.E.7

Results for LMME predicting mathematics achievement

	Estimate	SE*	t*	p*
(Intercept)	23.43	1.03	22.82	0.00
Prior Achievement	0.68	0.02	43.72	0.00
Basic Math	-1.38	0.48	-2.91	0.00
Other Math	-0.89	0.39	-2.25	0.02
Pre-Algebra	0.07	0.59	0.11	0.91
Algebra I	0.03	0.36	0.08	0.94
Geometry	0.54	0.46	1.17	0.24
Algebra II	2.16	0.40	5.41	0.00
Trigonometry	3.96	0.54	7.33	0.00
Other Advanced	0.87	0.34	2.58	0.01
Probability and Statistics	1.89	0.82	2.31	0.02
Other AP/IB Math	3.19	0.85	3.77	0.00
Pre-Calculus	5.96	0.51	11.74	0.00
Calculus	5.35	0.87	6.12	0.00
AP/IB Calculus	6.73	0.62	10.78	0.00
k_{Asian}	-0.37	0.19	-1.91	0.06
$k_{Black/African\ American}$	0.03	0.14	0.22	0.83
$k_{Hispanic}$	-0.06	0.10	-0.61	0.54
β^*_{Asian}	0.30	0.15	2.03	0.04
$\beta^*_{Black/African\ American}$	-0.07	0.10	-0.74	0.46
$\beta^*_{Hispanic}$	0.02	0.07	0.33	0.74
R^2		0.6067		
Adjusted R^2*		0.6047		
AIC*		45607.994		
BIC*		45739.835		

Note. *after accounting for weights and the design effect of 4

After allowing intercepts to differ by groups (LMMU), however, all the multiplicative constants ($k_g s$) and additive constants (β_g^*s) were statistically insignificant. This might be related to the model comparison results between LMME and LMMU ($F_{(6,3912)}$: 0.496, p = 1.000), such that allowing intercepts to vary for each focal group adds little information to LMME, so LMME is more parsimonious than LMMU.

Table 4.E.8

Results for LMMU predicting mathematics achievement

	Estimate	SE*	t*	p*
(Intercept)	24.31	1.29	18.79	0.00
Prior Achievement	0.66	0.02	32.63	0.00
Basic Math	-1.45	0.48	-3.01	0.00
Other Math	-0.94	0.40	-2.36	0.02
Pre-Algebra	0.02	0.60	0.04	0.97
Algebra I	-0.01	0.36	-0.04	0.97
Geometry	0.53	0.46	1.16	0.25
Algebra II	2.17	0.40	5.43	0.00
Trigonometry	4.00	0.54	7.37	0.00
Other Advanced	0.86	0.34	2.55	0.01
Probability and Statistics	1.90	0.82	2.31	0.02
Other AP/IB Math	3.24	0.85	3.82	0.00
Pre-Calculus	6.06	0.51	11.77	0.00
Calculus	5.43	0.88	6.19	0.00
AP/IB Calculus	6.85	0.63	10.82	0.00
Asian†	-0.89	4.81	-0.19	0.85
Black/African American†	-1.30	2.29	-0.56	0.57
Hispanic†	-2.46	2.05	-1.20	0.23
k_{Asian}	-0.39	0.21	-1.83	0.07

$k_{Black/African\ American}$	0.00	0.14	0.00	1.00
$k_{Hispanic}$	-0.12	0.11	-1.07	0.28
β^{*}_{Asian}	0.33	0.20	1.65	0.10
$\beta^{*}_{Black/African\ American}$	-0.03	0.13	-0.24	0.81
$\beta^{*}_{Hispanic}$	0.11	0.10	1.06	0.29
R^2		0.6068		
Adjusted R^2*		0.6045		
AIC*		45615.514		
BIC*		45766.188		

Note. *after accounting for weights and the design effect of 4; †the predictive mean difference (intercept) between this focal group and the reference (White) group

vi) FMM

FMM with a constant intercept (FMME) is used to model the situation when everyone has the same starting point, such as same average score in mathematics, but different slopes for each predictor between subgroups. A one-point increase in prior achievement is related to a .11 higher post achievement score for the Asian group than for the White group, which also suggests that the expected gain in mathematics achievement in the Asian group is larger than the White group.

The slopes of thirteen sequences of mathematics course-taking were statistically significant predictors of post-mathematics achievement of the White group, except for Pre-Algebra, Algebra I, Geometry, and Probability and Statistics. There was a significant difference in slopes (regression coefficients) between the Asian group and the White group in Algebra II (b = -4.44, p < .05) controlling for other variables. This suggests that the relationship between Algebra II and mathematics achievement differs between the

Asian group and the White group. That is, taking one more credit in Algebra II is negatively related to post mathematics achievement scores (*b* = 2.44-4.44= -2.00) for the Asian group, while the same event is positively related to post achievement scores (*b* = 2.44, *p* < .001) for the White group.

Table 4.E.9

Results for FMME predicting mathematics achievement

	Estimate	SE*	t*	p*
(Intercept)	23.38	1.03	22.64	0.00
Prior Achievement	0.67	0.02	38.14	0.00
Basic Math	-1.24	0.61	-2.03	0.04
Other Math	-1.11	0.51	-2.16	0.03
Pre-Algebra	0.02	0.77	0.03	0.97
Algebra I	-0.26	0.45	-0.57	0.57
Geometry	1.17	0.60	1.94	0.05
Algebra II	2.44	0.52	4.73	0.00
Trigonometry	3.74	0.67	5.61	0.00
Other Advanced	1.35	0.44	3.03	0.00
Probability and Statistics	1.96	1.02	1.92	0.06
Other AP/IB Math	3.45	1.06	3.25	0.00
Pre-Calculus	6.08	0.61	10.01	0.00
Calculus	4.94	1.01	4.89	0.00
AP/IB Calculus	6.40	0.74	8.64	0.00
Asian:Prior Achievement	0.11	0.05	2.15	0.03
Asian:Basic Math	-0.70	3.68	-0.19	0.85
Asian:Other Math	0.61	2.57	0.24	0.81
Asian:Pre-Algebra	2.33	5.16	0.45	0.65
Asian:Algebra I	-1.17	1.57	-0.74	0.46
Asian:Geometry	1.65	2.48	0.67	0.51
Asian:Algebra II	-4.44	2.15	-2.06	0.04
Asian:Trigonometry	-2.62	2.61	-1.00	0.32
Asian:Other Advanced	-0.74	1.78	-0.42	0.68
Asian:Probability and Statistics	-3.63	5.00	-0.73	0.47
Asian:Other AP/IB Math	-1.21	3.04	-0.40	0.69

Asian:Pre-Calculus	-3.30	2.22	-1.48	0.14
Asian:Calculus	-0.26	3.90	-0.07	0.95
Asian:AP/IB Calculus	-2.01	1.88	-1.07	0.29
Black:Prior Achievement	-0.05	0.03	-1.59	0.11
Black:Basic Math	-0.65	1.26	-0.52	0.61
Black:Other Math	0.55	0.89	0.62	0.54
Black:Pre-Algebra	0.11	1.46	0.07	0.94
Black:Algebra I	0.69	1.08	0.64	0.52
Black:Geometry	-1.36	1.23	-1.11	0.27
Black:Algebra II	0.33	1.18	0.28	0.78
Black:Trigonometry	-0.02	1.56	-0.01	0.99
Black:Other Advanced	-0.87	0.81	-1.07	0.28
Black:Probability and Statistics	2.11	2.37	0.89	0.37
Black:Other AP/IB Math	-2.65	3.26	-0.81	0.42
Black:Pre-Calculus	1.00	1.55	0.64	0.52
Black:Calculus	-2.22	3.50	-0.63	0.53
Black:AP/IB Calculus	1.13	2.58	0.44	0.66
Hispanic:Prior Achievement	0.00	0.02	0.12	0.90
Hispanic:Basic Math	0.11	1.08	0.10	0.92
Hispanic:Other Math	0.75	1.02	0.73	0.47
Hispanic:Pre-Algebra	-0.09	1.48	-0.06	0.95
Hispanic:Algebra I	1.03	0.75	1.37	0.17
Hispanic:Geometry	-1.93	1.06	-1.83	0.07
Hispanic:Algebra II	-0.74	0.92	-0.80	0.42
Hispanic:Trigonometry	1.54	1.46	1.06	0.29
Hispanic:Other Advanced	-1.40	0.91	-1.54	0.12
Hispanic:Probability and Statistics	-1.62	2.19	-0.74	0.46
Hispanic:Other AP/IB Math	-0.28	2.13	-0.13	0.89
Hispanic:Pre-Calculus	-0.66	1.25	-0.52	0.60
Hispanic:Calculus	3.08	2.64	1.17	0.24
Hispanic:AP/IB Calculus	1.06	1.66	0.64	0.53
R^2		0.6091		
Adjusted R^2*		0.6035		
AIC*		45692.106		
BIC*		46049.958		

Note. *after accounting for weights and the design effect of 4

When ethnic groups have different slopes and intercepts, that is the most complex model in this thesis, the full moderated model with varying (unequal) intercepts (FMMU). According to the results, the varying intercepts for each group were not significantly different from the intercept for the White group. This result was the opposite from the model without prior achievement. That is, prior achievement was different between groups, so that intercepts did not significantly differ from the reference group after accounting for prior achievement. Moreover, the relationship between prior achievement and post achievement for the focal group (Asian, Black/African American, and Hispanic) does not significantly differ from the White group, and the difference between the prior and post achievement for the Asian group became insignificant. Yet, there was significant differences in slopes (regression coefficients) between the Asian group and the White group in Algebra II (b = -4.77, $p < .05$) controlling for other variables. That is, taking one more credit in Algebra II is negatively related to post mathematics achievement scores (b = 2.40-4.77= -2.37) for the Asian group, while the same event is positively related to post achievement scores (b = 2.40, $p < .001$) for the White group. All other predictors except Algebra II for the Asian group do not statistically differ from those for the White group after accounting for the design effect.

Table 4.E.10

Results for FMMU predicting mathematics achievement

	Estimate	SE*	t*	p*
(Intercept)	24.00	1.42	16.89	0.00
Prior Achievement	0.66	0.02	32.50	0.00
Basic Math	-1.34	0.63	-2.13	0.03
Other Math	-1.17	0.52	-2.24	0.03
Pre-Algebra	-0.08	0.79	-0.10	0.92
Algebra I	-0.34	0.47	-0.73	0.47
Geometry	1.08	0.62	1.75	0.08
Algebra II	2.40	0.52	4.60	0.00
Trigonometry	3.71	0.67	5.56	0.00
Other Advanced	1.29	0.45	2.83	0.00
Probability and Statistics	1.95	1.02	1.91	0.06
Other AP/IB Math	3.46	1.06	3.26	0.00
Pre-Calculus	6.09	0.61	10.02	0.00
Calculus	4.97	1.01	4.91	0.00
AP/IB Calculus	6.43	0.74	8.67	0.00
Asian†	3.69	6.03	0.61	0.54
Black/African American†	-1.27	2.84	-0.45	0.66
Hispanic†	-1.92	2.46	-0.78	0.44
Asian:Prior Achievement	0.08	0.08	0.95	0.34
Asian:Basic Math	-1.27	3.79	-0.34	0.74
Asian:Other Math	0.36	2.61	0.14	0.89
Asian:Pre-Algebra	1.57	5.30	0.30	0.77
Asian:Algebra I	-1.40	1.64	-0.85	0.39
Asian:Geometry	1.05	2.66	0.40	0.69
Asian:Algebra II	-4.77	2.21	-2.15	0.03
Asian:Trigonometry	-2.93	2.66	-1.10	0.27
Asian:Other Advanced	-1.08	1.87	-0.58	0.56
Asian:Probability and Statistics	-3.61	4.99	-0.72	0.47
Asian:Other AP/IB Math	-1.05	3.05	-0.35	0.73
Asian:Pre-Calculus	-3.37	2.22	-1.51	0.13
Asian:Calculus	-0.13	3.93	-0.03	0.97
Asian:AP/IB Calculus	-1.90	1.89	-1.01	0.31
Black:Prior Achievement	-0.04	0.05	-0.81	0.42
Black:Basic Math	-0.50	1.30	-0.38	0.70
Black:Other Math	0.67	0.93	0.72	0.47

Black:Pre-Algebra	0.28	1.50	0.19	0.85
Black:Algebra I	0.91	1.20	0.76	0.45
Black:Geometry	-1.17	1.29	-0.91	0.36
Black:Algebra II	0.39	1.19	0.33	0.74
Black:Trigonometry	0.05	1.57	0.03	0.97
Black:Other Advanced	-0.74	0.85	-0.87	0.38
Black:Probability and Statistics	2.12	2.37	0.90	0.37
Black:Other AP/IB Math	-2.67	3.26	-0.82	0.41
Black:Pre-Calculus	0.97	1.55	0.62	0.53
Black:Calculus	-2.27	3.50	-0.65	0.52
Black:AP/IB Calculus	1.05	2.59	0.41	0.69
Hispanic:Prior Achievement	0.03	0.04	0.67	0.50
Hispanic:Basic Math	0.38	1.13	0.33	0.74
Hispanic:Other Math	0.93	1.05	0.89	0.38
Hispanic:Pre-Algebra	0.24	1.53	0.16	0.87
Hispanic:Algebra I	1.27	0.81	1.57	0.12
Hispanic:Geometry	-1.67	1.11	-1.51	0.13
Hispanic:Algebra II	-0.63	0.93	-0.67	0.50
Hispanic:Trigonometry	1.60	1.46	1.10	0.27
Hispanic:Other Advanced	-1.26	0.93	-1.35	0.18
Hispanic:Probability and Statistics	-1.65	2.19	-0.75	0.45
Hispanic:Other AP/IB Math	-0.34	2.13	-0.16	0.87
Hispanic:Pre-Calculus	-0.68	1.25	-0.54	0.59
Hispanic:Calculus	2.92	2.65	1.10	0.27
Hispanic:AP/IB Calculus	0.92	1.67	0.55	0.58
R^2		0.6092		
Adjusted R^2*		0.6033		
AIC*		45699.928		
BIC*		46076.614		

Note. *after accounting for weights and the design effect of 4; †the predictive mean difference (intercept) between this focal group and the reference (White) group

vii) Model Comparison

Model fit values are presented in table 4.E.11. In addition, all possible F-tests were performed to compare nested models using sum of squared errors and corresponding degrees of freedom for each model. The corresponding values and results are presented in table 4.E.12.

The ANCOVA model is significantly better than the base model ($F_{(3,3918)}$: 10.582, $p < .001$). Thus, there are group mean differences to consider, and one size does not fit all. The ANCOVA model (E2) has a higher adjusted R^2 value, and lower AIC and BIC values than the base model (E1).

PMME is significantly better than the base model ($F_{(3,3918)}$: 10.122, $p < .001$), which suggests that adding proportional slopes for each ethnic group helps to explain more variance in mathematics achievement. Comparing model fit values, PMME (D3) has a higher adjusted R^2 and lower AIC and BIC values than the base model (D1). When comparing PMMU with PMME, PMME is more parsimonious than PMMU ($F_{(3,3915)}$: .891, $p = 1.000$). Comparing ANCOVA and PMMU, it also turns out that ANCOVA is more parsimonious than PMMU ($F_{(3,3915)}$: .434, $p = 1.000$).

When comparing MMME with the base model, MMME is significantly better than the base model ($F_{(3,3918)}$: 10.455, $p < .001$) while MMME is more parsimonious than MMMU ($F_{(3,3915)}$: 0.441, $p = 1.000$). When comparing MMME with FMME or comparing MMMU with FMMU, MMME or MMMU was more parsimonious than the moderator models ($F_{(39,3879)}$=.716 and $F_{(39,3876)}$=.713, respectively, where all p = 1.000).

When comparing MMM with PMM, however, the degrees of freedom for both models are the same. In this case, any one of the model fit values can be used to compare models having the same degrees of freedom, since all favor the same model. MMME has a higher adjusted R^2 value, which means MMME is better than PMME considering the same degrees of freedom in both models. PMMU has a lower AIC value than MMMU, so PMMU is better than MMMU in this case.

LMME is more parsimonious than LMMU ($F_{(3,3912)}$: 0.496, p = 1.000), so allowing intercepts to vary for each focal group adds little information to LMME. On the other hand, LMMU is significantly better than both PMMU ($F_{(3,3923)}$: 1.257, p < .001) and MMMU ($F_{(3,3923)}$: 1.376, p < .001).

FMME was significantly better than both the base model ($F_{(42,3879)}$: 1.410, p < .001) and PMME ($F_{(39,3879)}$: 0.741, p < .001). However, MMME was more parsimonious than FMME ($F_{(39,3879)}$: 0.716, p = 1.000), and LMME was more parsimonious than FMME ($F_{(36,3879)}$: 0.666, p = 1.000). In addition, FMME was more parsimonious than FMMU ($F_{(3,3876)}$: 0.402, p = 1.000).

According to the results of all the *F*-tests for comparing between nested models, ANCOVA and LMME were found to be better than the other models. Although the ANCOVA model is not nested in LMME, an *F*-test could be performed to assess if the models differ significantly. LMME is significantly better than ANCOVA ($F_{(3,3915)}$: 1.195, p < .001). The adjusted R^2 favored LMME while AIC and BIC favored the ANCOVA model.

Table 4.E.11

Model fit values for each model

Model	R^2	**Adjusted** R^2	**AIC**	**BIC**
BASE	0.6031	0.6017	45625.334	45719.506
ANCOVA	0.6063	0.6046	**45602.583**	**45715.589**
PMME	0.6062	0.6045	45603.960	45716.966
PMMU	0.6065	0.6045	45610.289	45742.129
MMME	0.6063	0.6046	45602.964	45715.970
MMMU	0.6064	0.6044	45610.649	45742.489
LMME	0.6067	**0.6047**	45607.994	45739.835
LMMU	0.6068	0.6045	45615.514	45766.188
FMME	0.6091	0.6035	45692.106	46049.958
FMMU	**0.6092**	0.6033	45699.928	46076.614

Table 4.E.12

Results for F-tests of Model comparison

Models	Degrees of Freedom	*F*	*p*-value
Base vs. **ANCOVA**	(3, 3918)	10.582	<.001
Base vs. **PMME**	(3, 3918)	10.122	<.001
Base vs. **MMME**	(3, 3918)	10.455	<.001
Base vs. **LMME**	(6, 3915)	5.890	<.001
Base vs. **FMME**	(42, 3879)	1.410	<.001

ANCOVA vs. PMMU	(3, 3915)	0.434	1.000
ANCOVA vs. MMMU	(3, 3915)	0.315	1.000
ANCOVA vs. LMMU	(6, 3915)	0.846	1.000
ANCOVA vs. FMMU	(42, 3876)	0.684	1.000
PMME vs. PMMU	(3, 3915)	0.891	1.000
PMME vs. **LMME**	(3, 3915)	1.652	<.001
PMME vs. FMME	(39, 3879)	0.741	1.000
PMMU vs. **LMMU**	(3, 3912)	1.257	<.001
PMMU vs. FMMU	(39, 3876)	0.704	1.000
MMME vs. MMMU	(3, 3915)	0.441	1.000
MMME vs**. LMME**	(3, 3915)	1.322	<.001
MMME vs. FMME	(39, 3879)	0.716	1.000
MMMU vs. **LMMU**	(3, 3912)	1.376	<.001
MMMU vs. FMMU	(39, 3876)	0.713	1.000
LMME vs. LMMU	(3, 3912)	0.496	1.000
LMME vs. FMME	(36, 3879)	0.666	1.000
LMMU vs. FMMU	(36, 3876)	0.658	1.000
FMME vs. FMMU	(3, 3876)	0.402	1.000

Note: the bolded model is statistically better or more parsimonious model, effective *N*: 3,936 (after accounting for weights and the design effect of 4)

F. Ethnic differences in the relationship between prior achievement, *Level* and *Pattern* of mathematics course-taking and mathematics achievement

i) Base model

The base multiple regression model (Table 4.F.1) is the simplest model, and it assumes that all regression weights for all focal groups are equal to those of the reference group, so in other words, "one size fits all." For this model, we estimated one optimal intercept and one optimal slope coefficient per predictor. The results suggest that all predictors in this model (prior achievement, *Level*, and *Pattern*) were positively related to mathematics achievement, when controlling for the other variables in the model.

Table 4.F.1

Results for the base model predicting mathematics achievement

	Estimate	SE*	t*	p*
(Intercept)	23.01	1.00	23.01	0.00
Prior Achievement	0.68	0.02	44.82	0.00
Level	27.84	2.18	12.75	0.00
Pattern	5.85	0.28	20.61	0.00
R^2		0.6023		
Adjusted R^2*		0.6020		
AIC*		45600.805		
BIC*		45625.917		

Note. *after accounting for weights and the design effect of 4

ii) ANCOVA

The ANCOVA model assumes the same regression slopes for all groups but allows for different intercepts, which frees intercepts to vary from the base model. The results of ANCOVA show that all predictors as well as differences in intercepts for the Black and Hispanic groups were statistically significant (See table 4.F.2). This model assumes that the regression coefficients (slopes) of predictors are the same regardless of group membership, but the intercepts differ between ethnic groups. The means of the Black/African American and the Hispanic groups statistically differ from that of the reference group, the White group. The Black/African American and the Hispanic group of students, on average, had lower predicted mathematics scores (2.90 and 1.34 points lower, respectively) than the White group.

Table 4.F.2

Results for the ANCOVA model predicting mathematics achievement

	Estimate	SE*	t*	p*
(Intercept)	24.54	1.04	23.55	0.00
Prior Achievement	0.66	0.02	42.89	0.00
Level	28.09	2.18	12.86	0.00
Pattern	5.74	0.28	20.19	0.00
Asian†	1.50	1.00	1.50	0.13
Black/African American†	-2.90	0.58	-5.03	0.00
Hispanic†	-1.34	0.47	-2.88	0.00
R^2		0.6055		
Adjusted R^2*		0.6049		
AIC*		45577.812		
BIC*		45621.759		

Note. *after accounting for weights and the design effect of 4; †the predictive mean difference (intercept) between this focal group and the reference (White) group

iii) PMM

PMM constrains differences in slopes between the focal group and the reference group to be a multiplicative constant of the slope of the reference group. PMM with a constant intercept (PMME, Table 4.F.3) gives constraints on slope differences between groups without varying group intercepts. The vector of regression weights for the White group is multiplied by -0.07 to get the difference in vectors of regression weights (prior achievement, *Level*, and *Pattern*) between the Black/African American group and the White group. The vector of regression weights for the White group is multiplied by -0.03 to get the difference in vectors of regression weights (prior achievement, *Level*, and *Pattern*) between the Hispanic group and the White group.

When the multiplicative constant, k_g, is between -1 and 0, the regression coefficients for the focal group are (1+k_g) times the regression coefficients for the reference group, where (1+k_g) is greater than 0 and less than 1. Therefore, the increment in the expected mathematics achievement score is smaller for the Black/African American group and the Hispanic group than the White group. The absolute values of the regression coefficients for these focal groups are smaller compared to those for the regression coefficients for the reference group.

Table 4.F.3

Results for PMME predicting mathematics achievement

	Estimate	SE*	t*	p*
(Intercept)	23.72	1.01	23.58	0.00
Prior Achievement	0.68	0.02	44.32	0.00
Level	28.42	2.18	13.01	0.00
Pattern	5.78	0.29	20.26	0.00
k_{Asian}	0.02	0.02	1.03	0.30
$k_{Black/African\ American}$	-0.07	0.01	-4.97	0.00
$k_{Hispanic}$	-0.03	0.01	-2.71	0.01
R^2		0.6053		
Adjusted R^2*		0.6047		
AIC*		45579.454		
BIC*		45623.401		

Note. *after accounting for weights and the design effect of 4

All multiplicative constants, k_gs, in PMM with unequal intercepts (PMMU) were not statistically significant after controlling for differences in intercepts (Table 4.F.4). It is not surprising that PMME has lower AIC and BIC, since PMME is more parsimonious than PMMU. Additionally, differences in intercepts between a focal group and the reference group were not statistically significant compared to the White group in PMMU.

Table 4.F.4

Results for PMMU predicting mathematics achievement

	Estimate	SE*	t*	p*
(Intercept)	24.14	1.16	20.78	0.00
Prior Achievement	0.67	0.02	38.71	0.00
Level	28.30	2.20	12.85	0.00
Pattern	5.78	0.29	20.04	0.00
Asian†	4.47	3.48	1.28	0.20

Black/African American†	-1.50	1.79	-0.84	0.40
Hispanic†	-0.99	1.48	-0.67	0.50
k_{Asian}	-0.06	0.06	-0.90	0.37
$k_{Black/African\ American}$	-0.04	0.05	-0.81	0.42
$k_{Hispanic}$	-0.01	0.04	-0.23	0.82
R^2	0.6056			
Adjusted R^2*	0.6047			
AIC*	45585.480			
BIC*	45648.261			

Note. *after accounting for weights and the design effect of 4; †the predictive mean difference (intercept) between this focal group and the reference (White) group

iv) MMM

MMM constrains the difference in slopes between the focal group and the reference group to be a constant. When all regression coefficients between a focal group g and a reference group differ by an additive constant, β_g^*, the vectors of regression coefficients between those two groups differ only in their means. For MMM with a constant intercept (MMME, Table 4.F.5), the negative values of $\beta^*_{Black/African\ American}$ and $\beta^*_{Hispanic}$ suggest that a one-point increase in total predictor score leads to 0.06 and 0.02 point decreases, respectively in the expected mathematics achievement score compared to the expected mathematics achievement score for the White group.

Table 4.F.5

Results for MMME predicting mathematics achievement

	Estimate	SE*	t*	p*
(Intercept)	23.73	1.01	23.59	0.00
Prior Achievement	0.68	0.02	44.37	0.00

Level	28.02	2.18	12.83	0.00
Pattern	5.71	0.29	20.02	0.00
β^*_{Asian}	0.02	0.01	1.27	0.21
$\beta^*_{Black/African\ American}$	-0.06	0.01	-5.04	0.00
$\beta^*_{Hispanic}$	-0.02	0.01	-2.68	0.01
R^2		0.6055		
Adjusted R^2*		0.6049		
AIC*		45578.324		
BIC*		45622.271		

Note. *after accounting for weights and the design effect of 4

On the contrary, the results of MMM with unequal intercepts (MMMU, Table 4.F.6) show that all additive constants were no longer statistically significant after accounting for group differences in intercepts. Additionally, differences in intercepts for focal groups were not statistically significant compared to the White group in MMMU.

Table 4.F.6

Results for MMMU predicting mathematics achievement

	Estimate	SE*	t*	p*
(Intercept)	24.23	1.20	20.16	0.00
Prior Achievement	0.67	0.02	35.81	0.00
Level	28.05	2.18	12.84	0.00
Pattern	5.73	0.29	20.03	0.00
Asian†	2.66	4.03	0.66	0.51
Black/African American†	-1.18	1.98	-0.60	0.55
Hispanic†	-1.39	1.71	-0.81	0.42
β^*_{Asian}	-0.02	0.06	-0.30	0.76
$\beta^*_{Black/African\ American}$	-0.04	0.04	-0.91	0.36
$\beta^*_{Hispanic}$	0.00	0.03	0.05	0.96
R^2		0.6056		
Adjusted R^2*		0.6047		

AIC*	45585.871
BIC*	45648.652

Note. *after accounting for weights and the design effect of 4; †the predictive mean difference (intercept) between this focal group and the reference (White) group

v) LMM

LMM allows regression weight vectors to differ by both additive and multiplicative constants, so this model is a combination of PMM and MMM. For LMM with a constant intercept (LMME, Table 4.F.7), both the multiplicative constant (k_{Asian}) and the additive constant (β^*_{Asian}) were statistically significant for the Asian group of students. The difference in vectors of regression coefficients differs by -.45*β_{v0}+.38, where β_{v0} is the vector of regression coefficients of the White group of students. These additive or multiplicative constants were not significant for the Black/African American group and the Hispanic group.

Table 4.F.7

Results for LMME predicting mathematics achievement

	Estimate	SE*	t*	p*
(Intercept)	23.76	1.01	23.61	0.00
Prior Achievement	0.67	0.02	42.50	0.00
Level	28.96	2.42	11.97	0.00
Pattern	5.88	0.33	17.68	0.00
k_{Asian}	-0.45	0.20	-2.20	0.03
$k_{Black/African\ American}$	0.06	0.16	0.37	0.71
$k_{Hispanic}$	-0.07	0.11	-0.64	0.52
β^*_{Asian}	0.38	0.16	2.30	0.02

$\beta^*_{Black/African\ American}$	-0.10	0.12	-0.84	0.40
$\beta^*_{Hispanic}$	0.03	0.08	0.37	0.71
R^2		0.6060		
Adjusted R^2*		0.6051		
AIC*		45581.936		
BIC*		45644.717		

Note. *after accounting for weights and the design effect of 4

Furthermore, when intercepts were allowed to differ by group (LMMU, Table 4.F.8), groups did not differ in slopes except for the multiplicative constant for the Asian group. However, the additive constant was no longer significant after accounting for the design effect (b = .39, p = .05).

Table 4.F.8

Results for LMMU predicting mathematics achievement

	Estimate	SE*	t*	p*
(Intercept)	24.50	1.22	20.10	0.00
Prior Achievement	0.66	0.02	32.59	0.00
Level	29.36	2.45	11.98	0.00
Pattern	5.96	0.34	17.54	0.00
Asian†	-0.35	4.28	-0.08	0.94
Black/African American†	-1.30	2.03	-0.64	0.52
Hispanic†	-1.96	1.80	-1.09	0.28
k_{Asian}	-0.46	0.21	-2.12	0.03
$k_{Black/African\ American}$	0.03	0.16	0.21	0.83
$k_{Hispanic}$	-0.11	0.12	-0.95	0.34
β^*_{Asian}	0.39	0.20	1.94	0.05
$\beta^*_{Black/African\ American}$	-0.06	0.13	-0.44	0.66
$\beta^*_{Hispanic}$	0.10	0.10	0.93	0.36
R^2		0.6061		
Adjusted R^2*		0.6049		

AIC*	45589.631
BIC*	45671.246

Note. *after accounting for weights and the design effect of 4; †the predictive mean difference (intercept) between this focal group and the reference (White) group

vi) FMM

The full moderated model is the most complex model among all ten suggested models. FMM with a constant intercept (FMME, Table 4.F.9) is designed to model the situation when everyone has the same starting point, such as same average score in mathematics, but there are different slopes for each predictor between subgroups. A one-point increase in prior achievement is related to a 0.12 point higher in expected post achievement score for the Asian group than for the relationship between prior and post achievement scores for the White group.

Table 4.F.9

Results for FMME predicting mathematics achievement

	Estimate	SE*	t*	p*
(Intercept)	23.72	1.01	23.55	0.00
Prior Achievement	0.67	0.02	38.51	0.00
Level	30.09	2.87	10.49	0.00
Pattern	5.84	0.34	17.24	0.00
Asian:Prior Achievement	0.12	0.05	2.32	0.02
Asian:Level	-19.37	10.05	-1.93	0.05
Asian:Pattern	-2.01	1.09	-1.84	0.07
Black:Prior Achievement	-0.05	0.03	-1.68	0.09
Black:Level	0.05	5.89	0.01	0.99
Black:Pattern	0.31	0.84	0.37	0.71
Hispanic:Prior Achievement	0.00	0.02	-0.19	0.85

Hispanic:Level	-4.13	4.82	-0.86	0.39
Hispanic:Pattern	-0.29	0.61	-0.48	0.63
R^2		0.6061		
Adjusted R^2*		0.6049		
AIC*		45590.073		
BIC*		45671.689		

Note. *after accounting for weights and the design effect of 4

When ethnic groups have different intercepts in the full moderated model (FMMU, Table 4.F.10), that is the most complex model among all ten models. According to the results, the distinct intercepts for each group were not significantly different from the intercept for the White group. Moreover, the slope differences in prior achievement for the Asian group disappeared from FMME after accounting for other predictors in FMMU. Yet, there was significant difference in slopes (regression coefficients) for *Level* between the Asian group and the White group (b = -22.08, p = .047) controlling for other variables.

Table 4.F.10

Results for FMMU predicting mathematics achievement

	Estimate	SE*	t*	p*
(Intercept)	24.48	1.40	17.52	0.00
Prior Achievement	0.66	0.02	32.45	0.00
Level	29.40	3.00	9.80	0.00
Pattern	5.95	0.37	16.11	0.00
Asian†	3.42	5.74	0.60	0.55
Black/African American†	-1.47	2.72	-0.54	0.59
Hispanic†	-2.30	2.42	-0.95	0.34
Asian:Prior Achievement	0.08	0.08	0.99	0.32

Asian:Level	-22.08	11.10	-1.99	0.05
Asian:Pattern	-1.62	1.30	-1.25	0.21
Black:Prior Achievement	-0.04	0.05	-0.80	0.42
Black:Level	1.38	6.53	0.21	0.83
Black:Pattern	0.10	0.92	0.11	0.91
Hispanic:Prior Achievement	0.02	0.04	0.62	0.53
Hispanic:Level	-2.37	5.17	-0.46	0.65
Hispanic:Pattern	-0.65	0.72	-0.91	0.36
R^2		0.6062		
Adjusted R^2*		0.6047		
AIC*		45597.524		
BIC*		45697.974		

Note. *after accounting for weights and the design effect of 4; †the predictive mean difference (intercept) between this focal group and the reference (White) group

vii) Model Comparison

Goodness of fit values for each model are presented in table 4.F.11. In addition, *F*-tests were performed to compare nested models using sum of squared errors and corresponding degrees of freedom for each model. The corresponding *F*-test results are presented in table 4.F.12.

The ANCOVA model is significantly better than the base model ($F_{(3,3929)}$: 10.690, p < .001). There are group mean differences to consider, and one size does not fit all. The ANCOVA model (F2) has a higher adjusted R^2, and lower AIC and BIC values than the base model (F1). Also, having smaller absolute values of regression coefficients compared to the model without prior achievement is understandable since prior achievement partially mediates regression coefficients of other predictors.

PMME is significantly better than the base model ($F_{(3,3929)}$: 10.139, p < .001), which suggests that adding proportional slopes for each ethnic group helps to explain more variance in mathematics achievement. Comparing model fit, PMME (F3) has a higher adjusted R^2, and lower AIC and BIC values than the base model (F1). When comparing PMMU with PMME, PMME (the simpler model) is more parsimonious than PMMU ($F_{(3,3926)}$: .991, p = .612), and so allowing intercepts to vary in PMM model does not explain a significant amount of variance in mathematics achievement. Comparing ANCOVA and PMMU, it also turns out that the simpler model (ANCOVA) is more parsimonious ($F_{(3,3926)}$=.445, p = 1.000).

MMME is significantly better than the base model ($F_{(3,3929)}$: 10.518, p < .001), which suggests that adding additive slopes for each ethnic group helps to explain more variance in mathematics achievement. When comparing MMMU with MMME, MMME is more parsimonious than MMMU ($F_{(3,3926)}$: 0.485, p = 1.000). Comparing ANCOVA and MMMU, it also turns out that the simpler model (ANCOVA) is more parsimonious ($F_{(3,3926)}$: .315, p = 1.000).

LMME is significantly better than the base model ($F_{(6,3926)}$: 6.160, p < .001), which suggests that adding both proportional and additive slopes for each ethnic group helps to explain more variance in mathematics achievement than the base model. LMME is significantly better than PMME ($F_{(3,3926)}$: 2.171, p < .001) and MMME ($F_{(3,3926)}$:

1.795, $p < .001$), but MMME has the lowest AIC and BIC values among three models (PMME, MMME, and LMME) with constraints on slopes (See Table 4.F.11). When comparing LMMU with LMME, the simpler model (LMME) is more parsimonious ($F_{(3,3923)}$: 0.437, $p = 1.000$). Comparing ANCOVA and LMMU, it also turns out that ANCOVA is more parsimonious than LMMU ($F_{(6,3923)}$: 1.030, $p = .177$). On the other hand, LMMU is significantly different from both PMMU ($F_{(3,3923)}$: 1.615, $p < .001$) and MMMU ($F_{(3,3923)}$: 1.745, $p < .001$).

FMME was significantly better than both the base model ($F_{(9,3923)}$: 4.201, $p < .001$) and PMME ($F_{(6,3923)}$:1.230, $p < .001$). However, MMME was more parsimonious than FMME ($F_{(6,3923)}$: 1.042, $p = 0.099$), and LMME was more parsimonious than FMME ($F_{(3,3920)}$: 0.290, $p = 1.000$). In addition, FMME was more parsimonious than FMMU ($F_{(3,3920)}$: 0.518, $p = 1.000$).

According to the results of all the F-tests for comparing nested models, ANCOVA was found to be the most parsimonious model, but the assumptions for ANCOVA need to be checked. LMME was found to be better than the other nested models. The adjusted R^2 also favored LMME while AIC and BIC favored the ANCOVA model.

Table 4.F.11

Model fit values for each model

Model	R^2	**Adjusted** R^2	**AIC**	**BIC**
BASE	0.6023	0.6020	45600.805	45625.917
ANCOVA	0.6055	0.6049	**45577.812**	**45621.759**
PMME	0.6053	0.6047	45579.454	45623.401
PMMU	0.6056	0.6047	45585.480	45648.261
MMME	0.6055	0.6049	45578.324	45622.271
MMMU	0.6056	0.6047	45585.871	45648.652
LMME	0.6060	**0.6051**	45581.936	45644.717
LMMU	0.6061	0.6049	45589.631	45671.246
FMME	0.6061	0.6049	45590.073	45671.689
FMMU	0.6062	0.6047	45597.524	45697.974

Table 4.F.12

Results for F-tests of Model comparison

Models	Degrees of Freedom	F	p-value
Base vs. **ANCOVA**	(3, 3929)	10.690	<.001
Base vs. **PMME**	(3, 3929)	10.139	<.001
Base vs. **MMME**	(3, 3929)	10.518	<.001
Base vs. **LMME**	(6, 3926)	6.160	<.001
Base vs. **FMME**	(9, 3923)	4.201	<.001
ANCOVA vs. PMMU	(3, 3926)	0.445	1.000
ANCOVA vs. MMMU	(3, 3926)	0.315	1.000
ANCOVA vs. LMMU	(6, 3923)	1.030	0.177
ANCOVA vs. FMMU	(9, 3920)	0.810	1.000
PMME vs. PMMU	(3, 3926)	0.991	0.612
PMME vs. **LMME**	(3, 3926)	2.171	<.001
PMME vs. **FMME**	(6, 3923)	1.230	<.001
PMMU vs. **LMMU**	(3, 3923)	1.615	<.001
PMMU vs. FMMU	(6, 3920)	0.993	0.587
MMME vs. MMMU	(3, 3926)	0.485	1.000
MMME vs**. LMME**	(3, 3926)	1.795	<.001
MMME vs. FMME	(6, 3923)	1.042	0.099
MMMU vs. **LMMU**	(3, 3923)	1.745	<.001

MMMU vs. **FMMU**	(6, 3920)	1.058	0.039
LMME vs. LMMU	(3, 3923)	0.437	1.000
LMME vs. FMME	(3, 3923)	0.290	1.000
LMMU vs. FMMU	(3, 3920)	0.371	1.000
FMME vs. FMMU	(3, 3920)	0.518	1.000

Note: the bolded model is statistically better or more parsimonious model, effective *N*: 3,936 (after accounting for weights and the design effect of 4)

G. Testing the assumption for the ANCOVA model

Before performing ANCOVA, the assumption of the ANCOVA model (the same slopes for all groups), should be tested. The ethnic group accounts for a significant amount of variance in mathematics achievement (Table 4.G).

Table 4.G

Results for regression predicting mathematics achievement using ethnic group

	Estimate	SE*	t*	p*
(Intercept)	68.887	0.376	183.225	0.000
Asian	11.749	1.509	7.785	0.000
Black/African American	-11.782	0.854	-13.800	0.000
Hispanic	-7.224	0.702	-10.285	0.000
R^2		0.0811		
Adjusted R^2*		0.0799		
AIC*		28.7169		
BIC*		53.8286		

Note: *after accounting for weights and the design effect of 4

Secondly, the interaction between levels for the factor (ethnic group) and each predictor (one of thirteen sequences of course-taking or *Level* and *Pattern*) should be tested to see if the interaction is statistically significant and therefore needs to be included in the model. All interactions were not statistically significant for two models without prior achievement after accounting for the design effect (Tables 4.B.10 and 4.C.10). Therefore, the assumption of equal slopes is met.

However, adding prior achievement as a covariate, there was a significant interaction between Algebra II and mathematics achievement for the Asian group (p <

.05, Table 4.E.10) after accounting for the design effect, although it was just one out of fourteen interactions. There was also a significant interaction between *Level* and mathematics achievement for the Asian group ($p < .05$, Table 4.F.10) after accounting for the design effect. Therefore, the assumption for the ANCOVA model does not hold for E and F analyses with prior achievement.

V. Discussion

This chapter will cover 1) the best model among ten models comparing the results from all four types of analyses, 2) the effect of using *Level* and *Pattern*, 3) the effect of a covariate, prior achievement, and 4) limitation and suggestions from this study.

The best models among the ten models in the separate analyses were the Linearly Moderated Models (both LMMU and LMME). LMMU was the best when prior achievement was not included, while LMME was the best when accounting for prior achievement. When comparing ANCOVA and LMMU without prior achievement, LMMU was significantly better than ANCOVA ($F_{(6,3924)}$: 1.552 and $F_{(6,3913)}$: 1.677, respectively for analyses E and F, both $p < .001$). When accounting for prior achievement, LMME and ANCOVA were the two best models according to *F*-tests between nested models; however, the assumption for ANCOVA did not hold (Tables 4.E.10 and 4.F.10). In conclusion, when considering prior achievement, LMME was better than other nested models and it was more parsimonious than FMME. Meanwhile, without considering prior achievement, LMMU was more parsimonious than FMMU and it was significantly better than all other nested models.

LMMU in the B and C analyses suggested that all the multiplicative and additive constants for slopes were insignificant, but differences in intercepts were significant. Allowing intercepts to vary implies allowing differences in means between groups, and therefore the incremental effects of varying intercepts might be related to the decrease in significance of the additive constants in LMMU compared to LMME. The final two models were LMMU and ANCOVA; however, only intercepts (group means) differed between groups in LMMU although both the adjusted R^2 and F-tests favored LMMU over ANCOVA.

Accounting for prior achievement enables us to choose LMME in the E and F analyses, which may imply that prior achievement explains mean differences between groups. Both the multiplicative constant (k_{Asian}) and the additive constant (β^*_{Asian}) were statistically significant only for the Asian group in the E analysis, while only the additive constant was statistically significant for the Asian group (β^*_{Asian}) in the F analysis. The adjusted R^2 favored LMME over ANCOVA for Analyses C and D, and the assumption of ANCOVA did not hold while AIC and BIC favored the ANCOVA model over LMME.

It is possible that modeling the differences that exist between groups would be possible with only one of the three components: 1) allowing intercepts to vary, 2) adding an additive constant to the slope for the reference group, or 3) adding a multiplicative constant to the slope for the reference group. Beta weights for slopes or intercepts were not statistically significant although the Linearly Moderated Models were significantly better than other models. Also, the existence of multicollinearity among predictors, for example, between course-taking and additive constant to the slope, may be related to this conclusion regarding little significance in regression weights.

Another possibility for why none of the beta weights for slopes or intercepts were significantly better than other nested models is that the relationship between *Pattern* and mathematics achievement does not differ after accounting for *Level* for ethnic groups. Below are the unstandardized regression coefficients for Analyses C and F using *Level* and *Pattern* as predictors for each ethnic group and overall. Interestingly, unstandardized regression coefficients for intercepts and *Level* seem to differ between groups while those for *Pattern* and prior achievement seem to vary less between groups. This may suggest that all groups may have similar optimal patterns of course-taking, but their means (starting points) are different. However, the type III sum of squares for the *Pattern* were at least two times higher than those for *Level* for all models, which suggests that *Pattern*s explains more variation in mathematics achievement although the relationship between *Pattern* and achievement (unstandardized regression coefficient) does not differ substantially between groups.

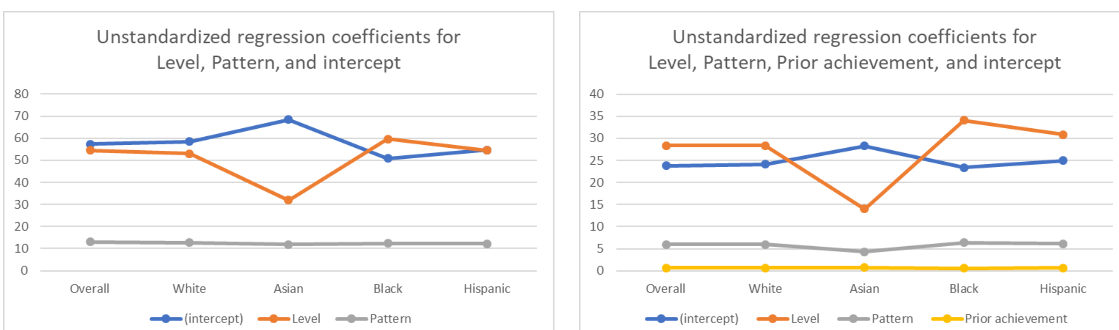


Figure 6. Unstandardized regression coefficients for the base model for predicting mathematics achievement with *Level* and *Pattern* (left) vs. Prior achievement, *Level*, and *Pattern* (right) for overall and each ethnic group

Comparing Level and Pattern vs thirteen sequences of course-taking (B vs C)

To state four analyses more simply, I would first compare the B analysis and the C analysis, neither of which used prior achievement when predicting mathematics achievement. The ten models in the B analysis have thirteen predictors, which are credits completed in each of thirteen sequences of mathematics course-taking. The ten models in the C analysis include *Level* and *Pattern* of credits taken in thirteen sequences of mathematics course-taking. Both analyses should have the same R^2 for the base model (B1 and C1) since the course sequences have been proven to be equivalent to *Level* and *Pattern* with respect to explaining the variance in the criterion (Davison & Davenport, 2002; Davison et al., 2015).

For the B analysis using thirteen predictors, LMMU was better than ANCOVA ($F_{(6,3913)}$: 1.677, p < .001) and more parsimonious than FMMU ($F_{(6,3880)}$: .858, p = 1.000). Meanwhile, in the C analysis using two predictors, *Level* and *Pattern,* LMM and FMM had the same degrees of freedom and the same values for all four measures of model fit (See Tables 4.C.8 and 4.C.10). For example, LMM and FMM in the C analysis, both LMMU and FMMU have the same total number of eleven predictors, but six of them are different with respect to their meanings. LMMU has one additive and one multiplicative constant for each focal group g, while FMMU has difference in Level and Pattern between each group g and the reference group. All of these six predictors were not statistically significant in both LMMU and FMMU in the C analysis. Interestingly, the number of predictors is the same in LMM and FMM in the C analysis while the number of predictors for LMM and FMM differs by 33 in the B analysis. For example, the model degrees of freedom for LMMU is 3,913 and for FMMU is 3,880 in the B analysis, while both LMMU and FMMU have degrees of freedom of 3,924 in the C analysis.

Effect of Prior achievement (Analyses B vs. E and C vs. F)

The direction of the slope for the Asian group's multiplicative constants in the two MMMU models differed depending on whether prior achievement was included when predicting mathematics achievement with thirteen predictors (B vs. E analyses). For example, the results for MMMU (Table 4.B.6) suggested that just increasing total predictor scores was negatively related to the criterion for the Asian group (b = -2.061, $p < .05$). The results showed that it was not helpful to consider increasing just the 'number' of courses taken for the Asian group of students when predicting mathematics achievement using the thirteen sequences, but the content of course mattered when prior achievement was not accounted for. However, after accounting for prior achievement, increasing total predictor scores for the Asian group increased the expected achievement by .302 points for MMMU (Table 4.E.6). In other words, increasing the sum or *Level* of courses also mattered for the Asian group after accounting for prior achievement, although it was not a huge effect.

In addition, differences in intercepts for the Asian group and the Black/African American group in MMMU became insignificant after accounting for prior achievement (Table 4.B.6 vs. Table 4.E.6). This suggests that prior achievement explains some variance in group means and contains some information about course-taking. These are Type III tests and prior achievement may be related to the other variables, which might be related to multicollinearity. The correlation between prior- and post achievement was .752 ($p < .001$) overall, and each group had different correlations between prior- and post achievement. The Asian group had the highest correlation between prior- and post achievement ($r = .796$, $p < .001$) for all ethnic groups, which seems to be related to the Asian group's additive constant for MMMU in Table 4.B.6 and Table 4.E.6. Compared to the same analyses when using *Level* and *Pattern* instead of the thirteen predictors (Tables 4.C.6 and 4.F.6), however, all differences in group means disappeared after accounting for prior achievement in MMMU. After accounting for prior achievement, the multiplicative constant for the Asian group ($b = -.418$, $p < .05$) in LMMU became significant in the analyses using *Level* and *Pattern*. All differences in group means disappeared after accounting for prior achievement in LMMU (Tables 4.C.8 and 4.F.8), which is the same phenomenon as MMMU. Therefore, prior achievement accounts for the mean differences between groups in most of the models suggested above.

Limitations and Suggestions

Results for all ten models did not differ substantially with respect to model comparison values (Tables 4.B.11, 4.C.11, 4.E.11, and 4.F.11), even though there were statistically significant differences between those models. For instance, the adjusted R^2 values for models with varying intercepts are the same (0.42) when rounded to the second decimal place (hundredths) in both B and C analyses. One the other hand, the adjusted R^2 values for models with a constant intercept are the same (0.60) in both E and F analyses, except for LMME model in the F analysis (0.61). Based on E and F analyses, MMME has the lowest values for AIC and BIC except for ANCOVA, where the assumption of equal regression slopes for each group is not met. The differences in AIC and BIC between PMME and MMME models in both E and F analyses differ for less than 1 in the analyses E and F, when the difference of 2 suggests weak evidence of a difference (Raftery, 1995). Therefore, much work remains to be done regarding the needs and effectiveness of presenting different models having different slopes and intercepts.

Less varying models could be caused by multicollinearity. Additive and multiplicative constants also had VIF higher than 10 as well as prior achievement and *Level* for groups in the full moderated model. This means group membership, prior achievement, and *Level* are highly correlated. Adding to this, the relationship between

Pattern of course-taking and achievement seems to be similar between groups after accounting for prior achievement.

In conclusion, the relationships of course-taking to achievement in mathematics were different between ethnic groups. Correlations (Table 4.A.3) were smaller for the Black/African American group and the Hispanic group, and low correlations could be because of low reliability. Therefore, course-taking (predictors) does not have the same meaning for minority ethnic groups compared to the White group. This could be because the achievement tests did not measure the same thing for different groups or because courses did not have the same meaning for different ethnic groups. The relationships between course-taking and achievement differ by ethnic group, which is related to different achievement gains between ethnic groups after controlling for prior achievement.

VI. Conclusion

This thesis introduced general linear model approaches to model the difference in relationships between predictors and the criterion using HSLS:09 data. Ten models, from one-size-fits-all to the full moderated multiple regression model, were introduced. Among these models, the Linearly Moderated Model (LMM) was statistically better than other simpler models, and it was more parsimonious than other more complex models when using HSLS:09 data to predict the relationship between mathematics course-taking and mathematics achievement. This application of the ten introduced models rendered few different results relative to one another, although there were statistically significant differences between models ($p < .05$). Yet, substantial insight about the relationship between course-taking and achievement can be drawn from this analysis of real data.

Interestingly, models allowing intercepts to vary (models with unequal intercepts) were significantly better than models with a constant (equal) intercept (e.g. LMMU was significantly better than LMME) when not accounting for prior achievement. On the other hand, the models with a constant intercept were more parsimonious than the models with unequal intercepts (e.g. LMME was more parsimonious than LMMU) when accounting for prior achievement as a covariate. Based on these results, this thesis suggests that prior achievement explains mean differences between different ethnic groups regardless of the model.

The relationships between course-taking and achievement in mathematics differ by ethnic group, however, only one component of the equation might be enough to model the differences that exist between groups: 1) allowing intercepts to vary, 2) adding an additive constant to the slope for the reference group, or 3) adding a multiplicative constant to the slope for the reference group. Beta weights for slopes or intercepts were not statistically significant although the Linearly Moderated Models were significantly better than other models. Also, the existence of multicollinearity among predictors, for example, between course-taking and additive constant to the slope, may be related to this conclusion regarding little significance in regression weights.

The ten suggested mathematical models also help inform and support our understanding of the achievement gap. Almost all models in this study showed that taking high-level courses such as Calculus are strongly and positively related to more gains in mathematics achievement for all students regardless of ethnicity. Taking higher-level mathematics courses in high school explains significant achievement gains, controlling for taking other courses and prior achievement for all ethnic groups, although the strength of the relationships between taking high-level courses and achievement differed between ethnic groups. The relationships between *Level* or *Pattern* of course-taking and achievement in mathematics also differed between ethnic groups, even though taking more courses as well as taking higher-level courses in mathematics explain significant variance in post mathematics achievement even after accounting for prior achievement for all ethnic groups.

Although prior achievement in high school students is highly related to the types of courses students choose to take, the relative benefit students obtain from courses may not vary depending on the level of prior achievement. Even though it might be the fact that all students have similar relationships between *Pattern*s of mathematics course-taking in high school and mathematics achievement, minority groups (here, Black/African American and Hispanic groups) had less optimal patterns compared to the White group. Also, mathematics tends to be hierarchical, for example, most students take Geometry and Algebra first and then Pre-Calculus and Calculus since students need some prior knowledge before taking high-level courses. Students who enter high school with different levels of prior achievement have different expected gains in achievement when taking the same courses due to the differences in their prior preparation coming to the courses. Although students tend to take difference courses, they may get different levels of benefit because they have different levels of prior achievement.

Moreover, studies have explored whether minority youth are given the same opportunities to reach advanced classes as their majority peers and found that an opportunity gap, which refers to unequal inputs (resources and opportunities), as compared to an achievement gap, which refers to unequal outputs (results and benefits), may also exist (Kelly, 2009; Riegle-Crumb & Grodsky, 2010). Students' participation in advanced courses in mathematics as well as the number of credits students completed in mathematics increased in all ethnic groups, but disparities in course-taking persisted in all subgroups including ethnic subgroups (Dalton et al., 2007). Also, research exploring longitudinal changes in achievement gaps has shown that achievement gaps have not disappeared (Shin, Davison, Long, Chan, & Heistad, 2013).

The achievement gap may exist before students enter elementary school or even earlier, although the gap is evident after students enter the school system. Once achievement gaps between student groups emerge, however, they tend to persist over time (Davison, Seo, Davenport, Butterbaugh, & Davison, 2004). High schools may be facing the task of overcoming already existing gaps in prior achievement. In addition, it is unrealistic and tougher for low-income and minority students to reduce the gap by learning faster than other more privileged students (Davison et al., 2004). These results imply that teachers or schools cannot be entirely blamed for the achievement gap, and therefore, it is not enough for high schools to prevent disadvantaged students from "falling behind" to narrow the achievement gap. There is no panacea for the achievement gap (Davison et al., 2004), and so there is no guarantee that just implementing a simple policy such as "Algebra in 8th grade" is effective to increase academic achievement and to close gaps.

Hence it is crucial that future research should focus more on early childhood education to lessen the gap when students are younger, but this does not mean that children must be prepared for standardized tests in early childhood. Early childhood education deserves more scholarly attention with respect to taking care of the undeserved. For example, initial mathematics skills and foundational concepts are important for later levels and allow for deeper understanding in mathematics (Claessens & Engel, 2013). Another study has shown that starting kindergarten with mathematics proficiency and experiencing a supportive home learning environment decreased SES-mathematics achievement gap significantly (Galindo & Sonnenschein, 2015). Playing numerical board games helped low-income children develop numerical knowledge and reduced the gap for preschoolers (Siegler & Ramani, 2008).

The achievement gap is not only limited to school education but is also related to bigger societal concerns and more general and systematic problems in human society. Educational inequality has been a challenge, and many efforts have been put in place to narrow achievement gaps. Still, the achievement gap should be explored more deeply to fully understand and solve ethnic disparities in achievement that exist beforehand. Ten suggested models in this study could also help answer questions we still have about the achievement gap in early childhood. Therefore, much work remains to be done utilizing these models and focusing on understanding and decreasing the achievement gap in students' earlier years.

References

Adelman, C. (1999). *Answers in the Tool Box. Academic Intensity, Attendance Patterns, and Bachelor's Degree Attainment.* Washington, DC: U.S. Department of Education. https://doi.org/ED431363

Adelman, C. (2004). *Principal Indicators of Student Academic Histories in Postsecondary Education, 1972-2000 U.S.* Washington, DC: U.S. Department of Education, Institute of Education Sciences. Retrieved from https://www2.ed.gov/rschstat/research/pubs/prinindicat/prinindicat.pdf

Aguinis, H. (2004). *Regression analysis for categorical moderators*. New York : Guilford Press.

Aiken, L. S., & West, S. G. (1991). *Multiple Regression: Testing and Interpreting Interactions*. SAGE publications, Inc.

Banas, P. A. (1964). *An investigation of transsituational moderators*. Unpublished Doctoral Dissertation. University of Minnesota.

Baron, R. M., & Kenny, D. A. (1986). The Moderator-Mediator Variable Distinction in Social Psychological Research : Conceptual , Strategic , and Statistical Considerations. *Journal of Personality and Social Psychology*, (6), 1173–1182.

Bohrnstedt, G., Kitmitto, S., Ogut, B., Sherman, D., & Chan, D. (2015). *School composition and the Black – White achievement gap*. (NCES 2015-018). U.S. Department of Education; Washington; DC: National Center for Education

Statistics. Retrieved from http://nces.ed.gov/pubsearch

Box, G. E. P., & Draper, N. R. (1987). *Empirical Model-Building and Response Surfaces*. John Wiley & Sons.

Bozick, R., & Ingels, S. J. (2008). *Mathematics Coursetaking and Achievement at the End of High School: Evidence From the Education Longitudinal Study of 2002 (ELS:2002)*. (NCES 2008-319). National Center for Education Statistics, Institute of Education Sciences, U.S. Department of Education. Washington, DC. Retrieved from papers3://publication/uuid/9B9F2FA3-9DC5-4E70-8CD2-BF6CDAA56DA8

Burkam, D. T., & Lee, V. E. (2003). *Mathematics, Foreign Language, and Science Coursetaking and the NELS:88 Transcript Data*. (NCES 2003–01), Washington, DC:U.S. Department of Education, National Center for Education Statistics.

Burnham, K. P., & Anderson, D. R. (2004). Multimodel Inference: Understanding AIC and BIC in Model Selection. *Sociological Methods & Research*, *33*(2), 261–304. https://doi.org/10.1177/0049124104268644

Cameron, C. E., Grimm, K. J., Steele, J. S., Castro-Schilo, L., & Grissmer, D. W. (2015). Nonlinear Gompertz curve models of achievement gaps in mathematics and reading. *Journal of Educational Psychology*, *107*(3), 789–804. https://doi.org/10.1037/edu0000009

Claessens, A., & Engel, M. (2013). How Important Is Where You Start? Early Mathematics Knowledge and Later School Success. *Teachers College Record*, *115*, 1–29.

Cohen, J., Cohen, P., West, S. G., & Aiken, L. S. (2003). *Applied Multiple Regression/Correlation Analysis for the Behavioral Sciences* (Third Edit). Mahwah, New Jersey.

Cook, R. D., & Weisberg, S. (2004). Partial One-Dimensional Regression Models. *The American Statistician*, *58*(2), 110–116. https://doi.org/10.1198/0003130043240

Dalton, B., Ingels, S. J., Downing, J., & Bozick, R. (2007). Advanced Mathematics and Science Coursetaking in the Spring High School Senior Classes of 1982, 1992, and 2004. *Center for Education Statistics Institute Of*. https://doi.org/10.1037/e608102011-001

Davenport, E. C., Davison, M. L., Kuang, H., Ding, S., Se-Kang, K., & Kwak, N. (1998). Mathematics Course-Taking by Gender and Ethnicity. *American Educational Research Journal*, *35*(3), 497–514.

Davenport, E. C., Davison, M. L., Wu, Y.-C., Kim, S.-K., Kuang, H., Kwak, N., … Ayodele, A. (2013). Number of courses, content of coursework, and prior achievement as related to ethnic achievement gaps in mathematics. *Journal of Educational Leadership in Action*, *2*(1). Retrieved from http://www.lindenwood.edu/academics/beyond-the-classroom/publications/journal-of-educational-leadership-in-action/all-issues/previous-issues/volume-2-issue-1/number-of-courses-content-of-coursework-and-prior-achievement/

Davenport, E. C., Park, K., Ayodele, A., Davison, M. L., & Kang, Y. (2015). Ethnic Differences in Mathematics Achievement Using Moderated Profile Analysis on ELS 2002 Data. Paper oral presented in 2015 annual meeting of the American Educational Research Association (AERA), Chicago, IL.

Davis-Kean, P. E., & Jager, J. (2014). Trajectories of achievement within race/ethnicity: "Catching up" in achievement across time. *Journal of Educational Research*, *107*(3), 197–208. https://doi.org/10.1080/00220671.2013.807493

Davison, M. L., & Davenport, E. C. (2002). Identifying criterion-related patterns of predictor scores using multiple regression. *Psychological Methods*, *7*(4), 468–484. https://doi.org/10.1037/1082-989X.7.4.468

Davison, M. L., Davenport, E. C., Chang, Y., Vue, K., & Su, S. (2015). Criterion-Related Validity : Assessing the Value of Subscores, *52*(3), 263–279.

Davison, M. L., Davenport, E. C., & Kohli, N. (2017). Population Differences in Linear Regression Coefficients: The Mean Moderated Model, Personal Communication.

Davison, M. L., Seo, Y. S., Davenport, E. C., Butterbaugh, D., & Davison, L. J. (2004). When Fall What Do Behind ? Can Be Children Done ? *The Phi Delta Kappan*, *85*(10), 752–761.

Forster, M. R. (2000). Key Concepts in Model Selection: Performance and Generalizability. *Journal of Mathematical Psychology*, *44*(1), 205–231. https://doi.org/10.1006/jmps.1999.1284

Friedrich, R. J. (1982). In Defense of Multiplicative Terms in Multiple Regression Equations, *26*(4), 797–833.

Galindo, C., & Sonnenschein, S. (2015). Decreasing the SES math achievement gap: Initial math proficiency and home learning environments. *Contemporary Educational Psychology*, *43*, 25–38. https://doi.org/10.1016/j.cedpsych.2015.08.003

Grooms, R. R., & Endler, N. S. (1960). The effect of anxiety on academic achievement. *Journal of Educational Psychology*, *51*(5), 299–304. https://doi.org/10.1037/h0042077

Hanushek, E. A. (1972). *Education and Race: An Analysis of the Educational Production Process*. Lexington, MA: D.C. Heath.

Hastedt, D. (2016). *Mathematics Achievement of Immigrant Students*. https://doi.org/10.1007/978-3-319-29311-0

Hoffer, T. B., Rasinski, K. A., & Moore, W. (1995). Social Background Differences in High School Mathematics and Science. *National Center for Education Statistics, Statistics in Brief*, (NCES 95-206).

Ingels, S. J., Pratt, D. J., Herget, D. R., Bryan, M., Fritch, L. B., Ottem, R., … Wilson, D. (2015). *High School Longitudinal Study of 2009 (HSLS:09) 2013 Update and High School Transcript Data File Documentation (NCES 2015-036).*

Ingels, S. J., Pratt, D. J., Herget, D. R., Burns, L. J., Dever, J. A., Ottem, R., … Leinwand, S. (2011). *High School Longitudinal Study of 2009 (HSLS:09). Base-*

Year Data File Documentation.

Ingels, S. J., Pratt, D. J., Herget, D. R., Dever, J. A., Fritch, L. B., Ottem, R., & Rogers, J. E. (2014). *High School Longitudinal Study of 2009 (HSLS:09) Base Year to First Follow-Up Data File Documentation (NCES 2014- 361)*.

Judd, C. M., McClelland, G. H., & Ryan, C. S. (2009). *Data analysis : a model comparison approach.* (2nd ed.). New York: Routledge.

Kam, C. D., & Franzese, R. J. (2007). *Modeling and Interpreting Interactive Hypotheses in Regression Analysis*. Ann Arbor : University of Michigan.

Kelly, S. (2009). The Black-White Gap in Mathematics Course Taking. *Sociology of Education*, *82*(1), 47–69. https://doi.org/10.1177/003804070908200103

Long, M. C., Conger, D., & Iatarola, P. (2012). Effects of High School Course-Taking on Secondary and Postsecondary Success. *American Educational Research Journal*, *49*(2), 285–322. https://doi.org/10.3102/0002831211431952

Marsh, L. C., & Cormier, D. R. (2002). *Quantitative applications in the social sciences: Spline regression models*. housand Oaks, Calif. ; London : SAGE. https://doi.org/10.4135/9781412985901

Peterson, E. R., Rubie-Davies, C., Osborne, D., & Sibley, C. (2016). Teachers' explicit expectations and implicit prejudiced attitudes to educational achievement: Relations with student achievement and the ethnic achievement gap. *Learning and Instruction*, *42*, 123–140. https://doi.org/10.1016/j.learninstruc.2016.01.010

Pitt, M. A., Myung, I. J., & Zhang, S. (2002). Toward a method of selecting among computational models of cognition. *Psychological Review*, *109*(3), 472–491. https://doi.org/10.1037/0033-295X.109.3.472

Raftery, A. E. (1995). Bayesian Model Selection in Social Research. *Sociological Methodology*, *25*, 111–163. Retrieved from http://www.jstor.org/stable/271063 Accessed : 29

Reardon, S. F. (2013). The Widening Income Achievement Gap. *Educational Leadership*, *70*(8), 10–16.

Reeves, E. B. (2012). The Effects of Opportunity to Learn, Family Socioeconomic Status, and Friends on the Rural Math Achievement Gap in High School. *American Behavioral Scientist*, *56*(7), 887–907. https://doi.org/10.1177/0002764212442357

Riegle-Crumb, C., & Grodsky, E. (2010). Racial-Ethnic Differences at the Intersection of Math Course-taking and Achievement. *Sociology of Education, 83*(3), 248–270. https://doi.org/10.1177/0038040710375689

Saunders, D. R. (1956). Moderator variables in prediction. *Educational and Psychological Measurement*, *16*(2), 209–222.

Schiller, K. S., & Muller, C. (2003). Raising the Bar and Equity ? Effects of State High School Graduation Requirements and Accountability Policies on Students ' Mathematics Course Taking. *Educational Evaluation and Policy Analysis*, *25*(3), 299–318.

Shin, T., Davison, M. L., Long, J. D., Chan, C. K., & Heistad, D. (2013). Exploring gains in reading and mathematics achievement among regular and exceptional students using growth curve modeling. *Learning and Individual Differences*, *23*(1), 92–100. https://doi.org/10.1016/j.lindif.2012.10.002

Siegler, R. S., & Ramani, G. B. (2008). Playing linear numerical board games promotes low-income children's numerical development. *Developmental Science*, *11*(5), 655–661. https://doi.org/10.1111/j.1467-7687.2008.00714.x

Smith, J. B. (1996). Does an Extra Year Make Any Difference ? The Impact of Early Access to Algebra on Long- Term Gains in Mathematics Attainment. *Educational Evaluation and Policy Analysis*, *18*(2), 141–153.

Son, J., Watanabe, T., & Lo, J.-J. (Eds.). (2017). *What Matters? Research Trends in International Comparative Studies in Mathematics Education*. https://doi.org/10.1007/978-3-319-51187-0

Stricker, L. J., Rock, D. a., & Burton, N. W. (1993). Sex differences in predictions of college grades from scholastic aptitude test scores. *Journal of Educational Psychology*, *85*(4), 710–718. https://doi.org/10.1037/0022-0663.85.4.710

Sturgis, P. (2004). Analysing Complex Survey Data: Clustering, Stratificaion and Weights. *Social Research Update*, *Autumn*(43).

Teitelbaum, P. (2003). The Influence of High School Graduation Requirement Policies in Mathematics and Science on Student Course-Taking Patterns and Achievement.

Educational Evaluation and Policy Analysisisluation, *25*(1), 31–57.

Vanneman, A., Hamilton, L., Baldwin Anderson, J., & Rahman, T. (2009). *Achievement Gaps: How Black and White Students in Public Schools Perform in Mathematics and Reading on the National Assessment of Educational Progress*. *Institute of Educational Sciences*. (NCES 2009-455). National Center for Education Statistics, Institute of Education Sciences, U.S. Department of Education. Washington, DC.

Yeh, S. S. (2017). Contradictions Resolved : An Analysis of Two Theories of the Achievement Gap. *Teachers College Record*, *119*(June), 1–42. Retrieved from http://www.tcrecord.org/content.asp?contentid=21786

Young, J. W. (1994). Differential prediction of college grades by gender and by ethnicity: a replication study. *Educational and Psychological Measurement*, *54*(4), 1022–1029.

Zedeck, S. (1971). Problems With the Use of “Moderator” Variables. *Psychological Bulletin*, *76*(4), 295–310. https://doi.org/10.1037/h0031543