

GED Student Achievement Trajectories and Predictors:
Using the Latent Class Growth Model to Understand Heterogeneous Learning Patterns

Lydia Fuller

Vanderbilt University

April 2014

Thesis completed in partial fulfillment of the Honors Program in Psychological Sciences under
the direction of Dr. Sonya Sterba

Abstract

Each year, 500,000 adults take the General Education Development (GED) exam, the gateway to a high school equivalency diploma. Nearly 40 percent fail. Little research exists on the relation between preparation experiences and exam success. This thesis investigated heterogeneous learning patterns among women at a GED preparation program in Detroit, Michigan during their first ten months of enrollment. Using Latent Class Growth Modeling, patterns among repeated measures on math and on reading were best accounted for by three latent trajectory classes per subject. Classes followed unique and non-overlapping achievement trajectories. Five factors were assessed as potential predictors of trajectory class membership. Younger age, non-English native language, and employment predicted a higher-achieving math trajectory, while being a native English speaker and being a single head of household predicted a higher-achieving reading trajectory. Last school grade completed was not predictive of achievement in either subject domain. Associations between math and reading achievement trajectories were also analyzed, with students generally exhibiting similar levels of achievement across subjects, though high math achievement trajectory membership was more predictive of high reading achievement trajectory membership than vice versa. Finally, membership in classes following higher achievement trajectories was determined to be associated with a greater likelihood of success on the GED exam.

GED Student Achievement Trajectories and Predictors:

Using the Latent Class Growth Model to Understand Heterogeneous Learning Patterns

Introduction

“These days, you need your education.”¹ “Without my education I’m not going to go far in life...to have a better life I need an education.” “We’re all women who stopped our education at young age”—and now, “we all here to get our GED...to better ourselves.”

These are the voices of women at Mercy Education Project, a nonprofit General Education Development (GED) exam preparation program for women in Detroit, Michigan. They speak of a reality that plagues populations across America: the gap between the academic credentials that an individual holds and those required for economic success.

The GED Certificate

Approximately 29.5 million American adults lack a high school diploma, representing 14.6 percent of the population over age 25. Of this group, 42 percent—12.4 million adults—never attended high school (U.S. Census Bureau, 2011). These individuals, however, may still enter higher categories of educational attainment and subsequently bolster their career and earnings potential by obtaining a “high school equivalent” credential. The most common of these is the GED certificate (Crissey & Bauman, 2012). In 2011, nearly 500,000 adults in the United States took the GED exam, and 62.5 percent (approximately 312,500) of these test-takers passed, thus earning their certificates. As of the end of 2012, approximately 19.5 million adults have passed the GED exam (GED Testing Service, 2012).

Passing the GED exam and earning a GED certificate is not only accompanied by transition into a higher “educational attainment” bracket on the U.S. Census, but also by

¹ Student quotations are derived from personal interviews conducted by the author in March 2012. Fourteen women total were interviewed, twelve of whom were also in the sample used for the quantitative analyses presented in this thesis.

enhanced career opportunities and a higher earnings potential. In the workforce, 96 percent of companies accept the GED certificate as a high school degree and over half of GED test takers cite employment opportunities as their motivation for taking the exam (GED Testing Service, 2012). In 2012, the unemployment rate among Americans with less than a high school diploma was 12.4 percent, versus 8.3 percent among high school graduates with no college education, 6.2 percent among Americans with Associate degrees, and 4.0 percent among those with Bachelor's degrees or higher (U.S. Bureau of Labor Statistics [BLS], 2012).

The wage differential between individuals with and without high school credentials has increased substantially over the past four decades (Heckman, Humphreys, and Mader, 2010). GED certificate holders now earn an average of \$3100 per month compared to \$2400 per month for adults with some high school education and \$2100 for adults with no high school education (Ewert, 2012). Though traditional high school diploma holders still have higher mean earnings, the figures suggest that obtaining a GED certificate raises projected earnings for adults with some or no formal high school education by 30 percent and 48 percent, respectively. Furthermore, nearly one in four GED certificate holders plan to enroll in a four-year college (GED Testing Service, 2012). Graduating from a four-year college would further raise individuals' average monthly earnings to \$4900, or \$58,800 per year (Ewert, 2012). The estimated lifetime earnings of an individual with high school credentials are \$260,000 higher than those of a high school dropout who does not return to school (Tyler & Lofstrom, 2009).

Earning a GED certificate, however, is not easy. It requires passing a series of tests in five subject areas (collectively the GED exam), a feat that an estimated 40 percent of current graduation-bound high school seniors would be unable to accomplish (Heckman et al., 2010). Individuals must pass tests in writing, reading, social studies, science, and mathematics, and the average score across these tests must exceed a minimum set by the administering jurisdiction

(typically an individual state). In most states, test takers must score at least 410 out of 800 on each subject, with an average score of 450 across the five tests (GED Testing Service, 2012).

GED Preparation Programs

GED preparation programs equip students with the knowledge and skills needed to pass the exam. Programs' formats, instructional methods, and staff and student profiles vary widely. For example, some programs emphasize one-on-one tutoring, while others focus on classroom instruction. Some place age, income, or other restrictions on participant eligibility, while others accept students from all backgrounds. Most programs are offered through not-for-profit agencies, but some of these rely heavily on public funds while others are sustained primarily by private donations. Programs also differ with regard to the academic level at which instruction is provided: while some offer resources for teaching adults basic literacy and computation skills, others presume all incoming students to be capable of coursework at the high school level (Ruzzi & Kraemer, 2006). Furthermore, some GED preparation programs were created as mandatory community corrections programs with attendance ordered and monitored by the courts, while others enroll students on a voluntary basis (Davis, Mottern & Ziegler, 2010).

In addition to broad variation across programs, students within programs vary widely on factors including age, last grade level completed, English language skills, and employment and parental statuses. Accommodating students' unique profiles and life situations has been identified as a crucial component of promoting student success (Ruzzi & Kraemer, 2006).

Current Empirical Understanding of GED Preparation Programs

Existing research related to GED preparation programs overwhelmingly focuses either on psychological and situational factors that affect students prior to enrollment or on post-graduation outcomes in higher education and the job market (e.g., Bujack, 2012; Cervero, 1983; Herring, 2012; Leininger & Kalil, 2011; Lewis, 2009). Wide research gaps exist concerning

students' experiences between enrolling in a GED preparation program and either graduating or dropping out. Specifically, there are gaps in knowledge pertaining to both the description and prediction of academic growth patterns that may vary across students.

1. Description of growth patterns. Current research fails to systematically address the wide variation found among GED students. GED student populations are treated as homogenous entities, with quantitative data focusing on demographics, retention, and GED attainment across entire programs. Even when researchers acknowledge some demographic and academic differences among students who enter GED programs, they neglect heterogeneity in students' individual experiences and academic growth patterns once enrolled. For example, Bujack (2012) and Leininger and Kalil (2011) explore self-efficacy and motivation among GED students at program entry as predictors of success, yet provide only program-wide correlations of these predictors with outcomes and do not explore ways in which these relationships may differ across students. Lewis (2009) reports that poor math and reading skills upon entry into GED programs adversely affect self-efficacy, but does not assess the impact of initial ability levels on subsequent within-program academic growth.

A variety of possible academic growth patterns, or learning trajectories, exist and may differ between students. First, students do not enter GED programs at identical levels of academic ability. Furthermore, initial levels may affect the speed of subsequent learning, and the form of such academic growth may vary greatly from one student to another. For instance, it is possible that students who start with higher competencies in a given domain are more prepared to learn new concepts and therefore accelerate faster. Alternately, some students starting with higher competency levels may become complacent with their present performance and experience lower levels of motivation, and thus accelerate more slowly. Additional variability may be found when multiple academic domains are considered. For example, students

progressing quickly in one domain (e.g., math) may grow more slowly in another domain (e.g., reading) if they are disproportionately focusing their efforts. For other students, progressing quickly in one domain may enhance transferrable academic skills or psychological factors such as self-esteem that in turn contribute to faster progress in other domains.

Thus, existing research not only fails to elucidate the complexity of student experiences between enrollment in GED preparation programs and graduation or dropping out, but also fails to elucidate the heterogeneity of these experiences across students.

2. Prediction of growth patterns. It is possible that different learning trajectories among GED students are predictable by certain pre-existing factors. For example, one early investigation conducted three decades ago found that program participants' highest grade completed was significantly related to performance on the GED test, with students who had completed more schooling faring better on the exam (Cervero, 1983). However, neither this nor subsequent studies delved into the learning course that bridges pre-program factors and post-program outcomes. Herring (2012) identifies three main categories of barriers to sustained participation in GED preparation programs: situational (i.e. lack of transportation and childcare; family not supporting continued education), institutional (i.e. program attendance requirements; tedious enrollment process), and dispositional (i.e. fear of being too old; hesitating to seem too ambitious; not enjoying studying), but did not quantify the actual effects of these barriers on students' academic growth once enrolled. Leininger and Kalil (2008) and Bujack (2012) highlight a positive correlation between students' self-efficacy and GED goal completion, but do not examine in-program learning growth patterns that might facilitate this relationship or define specific demographic factors that might be proxies for higher self-efficacy. A literature review conducted by Heckman and Rubinstein (2001) highlights the importance of motivation and self-discipline in determining GED students' success on the exam, but does not explore how these

skills might predict achievement patterns that foster the subsequent attainment of the GED certificate.

In sum, many descriptive and predictive questions remain unanswered about students' experiences within GED preparation programs. Little information is available to help educators understand GED student learning trajectories and predict certain subgroups' likelihood of being successful. Regarding description, what trajectories, or continuous patterns of change over time, do students follow between entering and existing GED preparation programs? How quickly do students obtain new skills and knowledge? To what extent does a student's learning growth in one subject area affect his/her learning growth in another subject area? Regarding prediction, are a student's pre-existing demographic factors related to his/her rate of academic growth within a given subject domain? How does this relate to ultimate success on the GED exam?

This thesis aims to contribute to educators', policymakers', GED-seekers', and other stakeholders' understanding of GED students and programs by examining heterogeneous learning trajectories and predictors of achievement at Mercy Education Project, a nonprofit adult education and GED preparation program for women in Detroit, Michigan.

Focal Population

As previously described, there is no "prototypical" GED preparation program, as programs vary widely in structure, method, and student population. However, evaluation of the student population at MEP is informative since MEP serves female students from a diverse range of ethnic backgrounds, family structures, and academic attainment levels (Mercy Education Project [MEP], 2012). Furthermore, MEP's exclusive focus on women lends it to more detailed investigation of an understudied subgroup in the GED student population. Women represent 42 percent of GED test takers in the U.S., but male test takers are disproportionately more likely to pass the exam (GED Testing Service, 2012).

Deeper analysis into the unique circumstances and academic experiences of female GED students is valuable in understanding a significant subgroup that may vary from their male counterparts. One analysis of high school dropouts found females to be significantly more likely than males to report “family constraints” as a barrier to sustained participation in a GED preparation program, and also more likely to name “perceived inability” and “logistical barriers” as reasons for not entering and persisting in programs (King, 2002). In fact, many students at MEP reported gender-related challenges in seeking to further their education. One student dropped out of twelfth grade when she gave birth to her daughter. Another dropped out after her freshman year of high school to care for seven children when their mother, her sister, was murdered. Several other students reported attending MEP in secret, as they feared that their husbands would not approve of them taking time to further their educations.²

Enrollment at MEP is voluntary, though daily attendance (Monday through Thursday) is required for all students. Current students range in age from 20 to 48. Forty-three percent are African American, 29 percent are Hispanic, 9 percent are White, 5 percent are Arabic, and 14 percent are multiracial or other (MEP, 2013).

MEP accepts new students four times per year, and classes from September thru July. In 2012-13, MEP served a total of 170 women (MEP, 2013). Though all students seek a GED certificate, they are not required to enter at the “GED Preparation” level (approximately eleventh grade). Given their wide range of entrance levels and rates of academic growth, the amount of time that students spend in the program before either passing the GED or dropping out varies widely (MEP staff, personal communications, 2012-2013).

Students attend MEP from 9 a.m. to 1 p.m., Monday through Thursday, for a four-hour

² This information is derived from personal interviews conducted by the author in March 2012. Fourteen women total were interviewed, twelve of whom were also in the sample used for the quantitative analyses presented in this thesis.

day that includes independent studying, one-on-one tutoring, and classroom instruction. There is no program fee, and MEP covers GED testing costs for students. Free lunch, transportation, and childcare are provided (MEP, 2012). There is no income cutoff for enrollment eligibility, but MEP students fall unanimously in the lowest income category indicated on self-report entrance forms, defined in terms of income by family size and ranging from \$14,900 for a single person to \$28,100 for a family of eight.

Thirty-eight MEP students earned a GED certificate in 2013 (MEP staff, personal communication, January 2014). In a typical year, eighty percent of students who participate for at least two months improve their academic skills by at least one grade level, and 70 percent of students improve by two or more grade levels (MEP, 2012). Such sample-wide figures, however, fail to elucidate which women are improving, which are not, and why. These figures fail to reveal whether certain subsets of the female student population might start at different academic skill levels and grow at different rates; they also do not illuminate what factors predict experiencing academic gains in the program.

Latent Class Growth Model

To accommodate the heterogeneity in learning anticipated within the population of GED students at MEP and explore whether this results in identifiable subsets of learning trajectories among students, this analysis employed a form of mixture modeling called Latent Class Growth Modeling (LCGM; also known as the groups-based trajectory model; Nagin, 1999). Mixture modeling in general is a set of statistical strategies that can be used to identify unobserved population subgroups, or “classes,” having different patterns of outcome scores and/or different predictor-outcome relationships (McLachlan & Peel, 2000). Traditional regression analysis accommodates the use of predictors such as age, gender, or ethnicity to explain variance in a *single outcome* for the population *as a whole*, using cross-sectional data. Mixture modeling is an

alternative statistical technique that can account for associations among *multiple* outcomes (measured cross-sectionally or longitudinally) using “latent classes” of individuals who vary on their mean level of these outcomes. Furthermore, mixture modeling accommodates complex predictor-outcome relationships by allowing the effect of each predictor (e.g., age) to differ across subgroups that are defined by, for instance, unique mean patterns of change in outcomes over time. By contrast, traditional regression analysis would require the explicit incorporation of product terms (e.g., age*time) in order to allow the effects of one predictor to differ across levels of another in the population (Sterba & Bauer, in press).

LCGM, specifically, identifies latent classes of individuals who vary on their patterns of change across a repeatedly measured construct. Each individual in the population being analyzed is assumed to adhere to a particular class-specific trajectory of quantitative change over time in his/her scores on this construct. However, the process that generates individual differences in these trajectories is unobserved. By separating the overall population into classes that follow similar trajectories, researchers may clearly and parsimoniously assess, as well as predict, heterogeneous patterns in the underlying processes of change.

For this analysis, scores on academic assessments administered throughout students’ time at MEP on two key subjects, math and reading, served as repeated measures. Math and reading outcomes were chosen in particular due to the central role that ability in these subjects plays in future career and financial success (i.e., Richie & Bates, 2013). Additionally, administrators at MEP reported that they repeatedly find the math sections (math computation and applied math) and reading section on the repeated assessments to be most predictive of students’ ultimate success on the GED exam (MEP staff, personal communication, November 2012).

Predictor Variables and Distal Outcomes

LCGM allows researchers to add covariates, or predictor variables, to the model to assess

which factors increase an individual's likelihood of following a particular change trajectory. As previously discussed, little research exists regarding students' in-program learning trajectories, including potential predictors of such patterns. However, two predictors—last grade completed (lower attainment) and age at program entry (being older)—have specifically been identified as potential barriers to initial enrollment and persistence in GED preparation programs (Cervero, 1983; Herring, 2012). This thesis investigated whether they also affect in-program learning growth over time. Additionally, Herring (2012) and King (2002) cite time commitments and family obligations as barriers to enrollment and persistence, and this thesis explored the potential effect of these factors through more specific proxies: employment status at program entry and whether a student is a female single head of household (FSHH; raising children without another adult present to burden-share). A final predictor variable, speaking a language other than English at home, was also examined. This variable was hypothesized to be associated with lower initial achievement levels in both math and reading as well as slower improvement. Though not the focus of prior empirical research, speaking a language other than English at home might lead to added difficulties in receiving instruction and taking tests in English.

In total, then, five predictor variables were assessed in this thesis to determine whether they significantly affect individuals' trajectory class memberships: last grade completed in school, age at program entry, language spoken at home, employment status at program entry, and whether the student was a single head of household.

In addition to accommodating potential predictor variables, LCGM allows the incorporation of "distal outcomes" into the model, by which researchers may analyze individuals' likelihood, based on class membership, of experiencing certain outcomes in the future (after the time at which the final observed measurement is collected; Muthén, 2004). A distal outcome of key importance in this study was whether students obtained GED certificates.

Research Questions

Using the LCGM techniques described above, this analysis addressed four key questions regarding student achievement in GED preparation programs, particularly at MEP:

1. For math and reading outcomes considered separately: can variations in student growth be explained by a parsimonious number of homogeneous latent classes existing within the overall population? If so, what number of classes and functional form (i.e., trajectory shape) of each class best accommodates the population's heterogeneity in math and in reading?
2. Can students' trajectory class memberships in math and in reading be predicted by pre-existing demographic factors, specifically last grade completed in school, age at program entry, language spoken at home, employment status at program entry, and whether the student is a single head of household?
3. Do students' achievement growth trajectories on math predict students' achievement growth trajectories on reading, and vice versa?
4. Are students' trajectory class memberships in math and in reading predictive of future success in obtaining a GED certificate?

Additionally, the quantitative LCGM analysis was supplemented with qualitative data from personal interviews with MEP students. This data provided contextual detail to the analysis and assisted in interpreting quantitative results.

Method

Sample

Data for this thesis was drawn from the student population at Mercy Education Project (MEP) in Detroit, Michigan, as previously discussed. The study received approval from the Vanderbilt University Institutional Review Board (#131606). The thesis sample included 100

MEP students. Due to the rolling nature of entrance and exit, these students did not all attend MEP simultaneously. Students were selected on the basis of having first enrolled in MEP between 2004 and 2012 and having scores on all five predictor variables as well as at least one repeated measure score on both the math and reading growth assessments. Women in the sample varied widely in background characteristics, incoming ability levels, and amount of time spent at MEP. As of June 2013, 55 of the women had taken and passed all five sections of the GED exam, 33 had left MEP without earning a GED certificate, and twelve were still attending the program. According to self-report entrance forms, the average age at which the women entered MEP's program was 32.5 (range 18-77), and the average grade attainment prior to dropping out of high school was 9.5 (halfway through ninth grade; range 5-11). Ninety-four percent of the sample was non-white: 62 percent African American, 16 percent Hispanic, 9 percent Arabic, 3 percent African, 2 percent Jamaican, and 2 percent mixed race. Eight percent did not speak English at home, 50 percent were single heads of household (raising children without an adult male present), and 89 percent were unemployed.

Measures

The observed outcome measures in this analysis were MEP students' scores on repeatedly administered Tests of Adult Basic Education (TABE). TABE is a set of assessments in math computation, applied math, reading, English language, vocabulary, language mechanics, and spelling. It is designed and widely used to measure curriculum mastery and skill achievement in adult basic education programs (Bujack, 2012). TABE's issuer, McGraw Hill, touts it as "the most comprehensive and reliable academic assessment product in adult basic education" (McGraw Hill, 2013). Indeed, one study comparing multiple longitudinal assessments found TABE scores to be the most predictive of students' ultimate GED test scores (Moore & Davies, 1984).

All students in the sample were tested with TABE upon entry into the program, and subsequent tests were administered repeatedly to assess learning gains in each subject. TABE tests were generally administered after every 120 hours of instruction that a woman received. Due to students' varying attendance records and/or students taking extended periods of time away from the program, temporal intervals between TABE test administrations were not necessarily identical within or between subjects.

As mentioned previously, for purposes of this analysis student achievement throughout the program was represented by two sets of longitudinal data: math scores over time (each an average of the two math section scores, math computation and applied math, on a given testing date) and reading scores over time. Possible scores on any single section of TABE ranged from 0 to 12.9, the latter denoting high school graduation level (12 years and 9 months of schooling).

The covariates, or predictor variables—last grade completed in school, age at program entry, language spoken at home, employment status at program entry, and whether the student was a single head of household—were measured for each student at program entry based on self-report enrollment forms. Since LCGM does not allow for missing covariate data due to the fact that the model places assumptions on the conditional distribution of outcomes given covariates (Nagin, 2005), any incomplete enrollment forms were supplemented by conversations with MEP employees who were aware of the students' statuses in the covariate categories. The first two covariates were treated as continuous variables, and the latter three were treated as binary variables. The five covariates were denoted in the model as $X_1 - X_5$, where:

- X_1 : Last grade completed in school; continuous
- X_2 : Age at program entry; continuous
- X_3 : Language spoken at home; binary (0=English, 1=other)
- X_4 : Employment status at program entry; binary (0=unemployed; 1=employed)
- X_5 : Single head of household; binary (0=no; 1=yes)

Design and Missing Data Handling

Given the individualized nature of instruction at MEP, TABE testing dates were not standardized across students, nor was the time span between two given tests. Although there are alternative ways of handling individual variation in measurement occasions (see Sterba, in press), rounding and binning approaches are common when LCGMs are fit as multivariate models (wide-format data) to avoid estimation problems involving time points where few persons contribute observations. Here a binning approach was used involving coding test dates by number of months since student entry, and then creating time “bins” such that T_0 represents test scores from the entrance exam; T_1 , 1-2 months post-entry, T_2 , 3-4 months post-entry; T_3 , 5-6 months post-entry; T_4 , 7-8 months post-entry; and T_5 , 9-10 months post-entry. TABE scores from after ten months post-entry were not included due to extreme sparsity, as most students in the sample had either passed the GED exam or dropped out of the program by this point. As MEP programs are suspended for two months each summer, the ten-month time frame used for this analysis represents one “school year” in the program.

No students had multiple testing points within a single time bin, though many lacked scores within some bins. This missing data was assumed to be Missing at Random (MAR) using the terminology of Rubin (1976), which means that intermittent missingness and/or dropout was due to observed variables in the model and not to unobserved variables (e.g., Muthén & Shedden, 1999). MAR allows for the probability of missingness at a given time point to be related to observed TABE scores from other time points and/or any of the observed covariates. For example, under MAR, TABE scores might be lacking for a student at T_4 due to her having scored highly at T_3 and thus having passed the GED exam prior to T_4 . Alternately, under MAR, a missing T_4 score could be due to a student having had a low observed score at T_3 such that she became discouraged and dropped out of the program, or could be due to the student being a

single head of household and needing to temporarily drop out to care for a sick child.

Analysis Procedure

1. Unconditional LCGM. The first step in this analysis was to fit students' repeated TABE scores on a given outcome (math or reading) with an unconditional LCGM model (no covariates predicting class membership) with K latent classes. Classes were denoted $k=1\dots K$. Each class was defined by a learning trajectory with class-specific growth parameters. Although LCGM allows class membership to be probabilistic (each student has a nonzero probability of being in each trajectory class), students can be "assigned" to their most likely class, as will be explained subsequently. One unconditional LCGM model was fit for math, and a separate model for reading, such that a student's class assignment for one subject had no impact on her class assignment for the other. This analysis and subsequent analyses in this study were conducted using *Mplus*, a statistical software program designed to accommodate mixture models (Muthén & Muthén, 1998-2013).

The k th class trajectory can be expressed by a regression equation in which the predicted outcome score (\hat{y}_{it}) for person i at time point t in the program is a function of time (index $i=1\dots N$ and $t=1\dots T$). For example, if the k th class-specific trajectory is linear, the equation would be: $\hat{y}_{it} = n_0^{(k)} + n_1^{(k)} \text{time}_{it}$, where $n_0^{(k)}$ is the class-specific intercept and $n_1^{(k)}$ is the class-specific slope. k superscripts denote that these intercept and slope growth coefficients can differ across classes.

Within a given class, LCGM assumes that individual scores at a given time will vary only due to random deviation, rather than due to systematic heterogeneity among members of the class. An individual's (i) score at any given time point (t) is therefore a function of both class-specific growth parameters and time-specific individual random deviation (ϵ_{it}) from the class trajectory, as follows: $y_{it} = n_0^{(k)} + n_1^{(k)} \text{time}_{it} + \epsilon_{it}$. Time-specific individual deviations are assumed to be uncorrelated within class (known as the assumption of local independence) and normally

distributed within class k , $\varepsilon_{it} \sim N(0, \sigma_t^{2(k)})$, where $\sigma_t^{2(k)}$ is the residual outcome variance for the k th class at time t . These residual variances were, as is typical, constrained equal across class (Nagin, 2005), implying that $\sigma_t^{2(k)} = \sigma_t^2$.

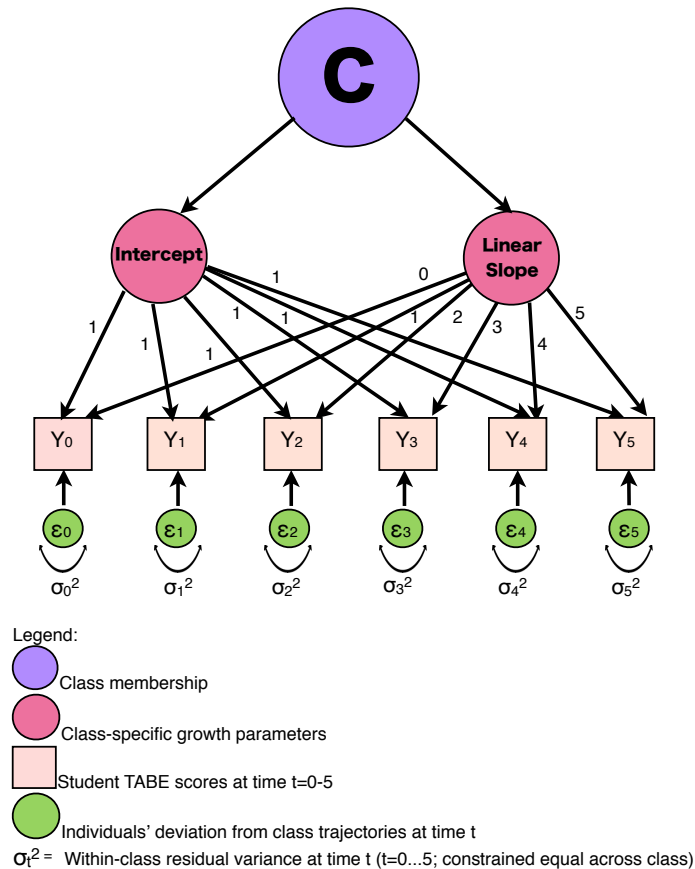
The unconditional LCGM also estimates the proportions of the sample in each class based on students' observed learning trajectories. These class proportions are estimated using a multinomial logistic regression specification:

$$p(c_i = k) = \exp(\omega^{(k)}) / \sum_{k=1}^K \exp(\omega^{(k)}) \text{ where } \omega^{(K)} = 0$$

In the above equation, $p(c_i = k)$ refers to the population-wide probability of membership in class k , such that class membership probabilities range from 0 to 1 for each class and sum to 1 across classes. ω^k refers to a multinomial intercept for class k . In order to avoid redundancy, only $K-1$ multinomial intercepts are estimated and then used to solve for all K class probabilities. The final ω^K is fixed at 0.

The K -class unconditional linear LCGM fit here has $q=(K-1) + T + (Kd)$ estimated parameters: one class proportion for every class but the last class, one residual variance for each time point (constrained to be constant across classes) and d growth coefficients (intercept, slope, etc.) per class, where here the number of growth coefficients, or polynomial order, was held constant across class.

An unconditional LCGM in which each class follows a linear trajectory is illustrated by the following path diagram:



In this study, LCGM thus allowed for an analysis of the underlying structure of heterogeneous classes of learners within MEP's student population as well as the prevalence of various learning patterns among students.

2. Selecting number and form of classes. The optimal number and optimal functional form (here, polynomial order d) of classes in a population are not determined within a single fitted LCGM. Rather, K and d are specified prior to running a given model. Model selection can be used to compare the fit of LCGMs differing in the number of classes and/or the functional form of each class trajectory (e.g., linear, quadratic, cubic). In line with research methods used by Nagin (1999), the number of classes that best accommodate the data was determined using a flexible nonlinear functional form. Quadratic growth was determined to be a generous starting allowance for form complexity after inspecting the functional form evidenced by plots of observed individual trajectories. After the number of classes was selected, more parsimonious

functional forms were tested and the “best fitting” model was chosen.

In mixture modeling, the “best fit” is determined using model selection indices. Indices known as Information Criteria are commonly used for selecting the number and form of classes, including Bayesian Information Criteria (BIC), sample-size adjusted BIC, and Akaike Information Criteria (AIC). Of these, BIC is most commonly used and has performed most consistently well in simulation evaluations for recovering the true number of classes when they indeed exist (Nylund, Asparouhov & Muthén, 2007; Tofighi & Enders, 2007). BIC was therefore the determining criteria for this analysis. The number and form of classes resulting in the lowest BIC was selected. Graphical inspection of plots of observed individual trajectories suggested that functional forms did not differ across classes, so the polynomial order was constrained to be equal across classes in order to maximize parsimony.

This model fitting process was conducted separately for each set of repeated measure scores. Thus, there were a certain number of classes and certain functional form for student reading growth, and a certain number of classes and certain functional form for student math growth.

Following identification of the optimal number and form of classes, LCGM was used to assign individuals to classes using “modal class assignment.” This technique involves first calculating the probability of an individual being in each class given her pattern of observed scores across time and model estimates. Next, each individual is “assigned” to the class for which she has the highest posterior probability of membership. These probabilities are referred to as “posterior probabilities” because they are calculated using Bayes' rule after the K class-specific trajectories have been estimated.

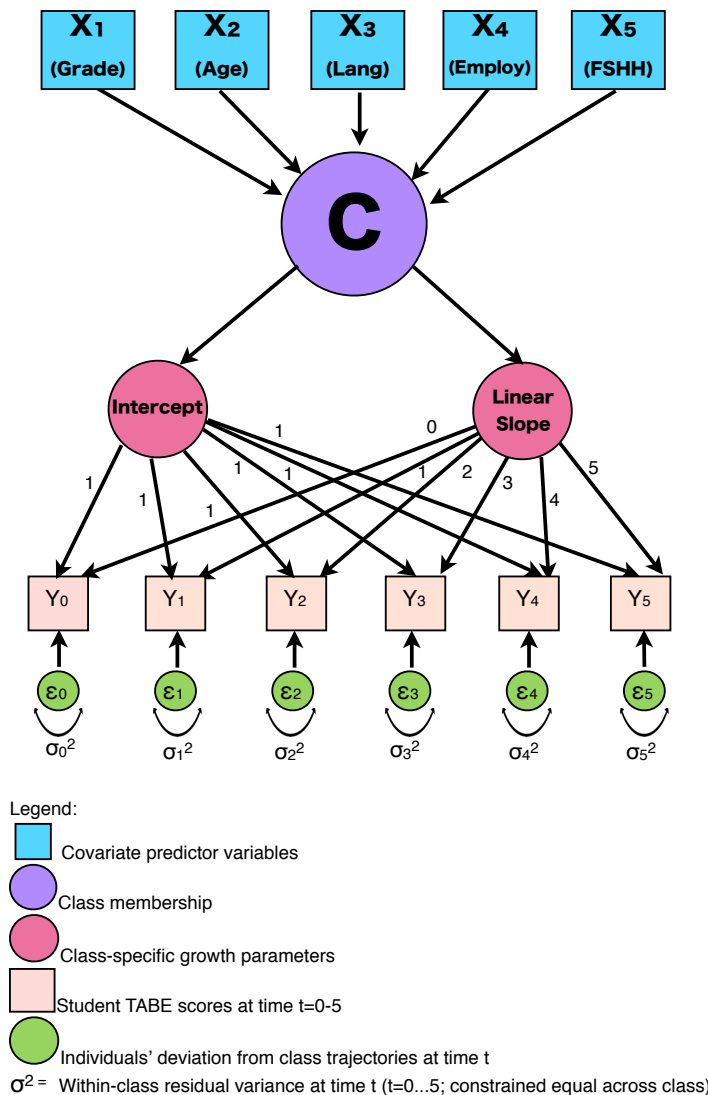
After all individuals were assigned, posterior probabilities were averaged across individuals assigned to each class, elucidating the overall accuracy of classification. For

example, if, in a model where $K=2$, individuals assigned to Class 1 had on average a .60 posterior probability of membership in Class 1 and on average a .40 posterior probability of membership in Class 2, classification accuracy would be considered quite poor. On the other hand, if they had on average a .98 posterior probability of membership in Class 1, classification accuracy would be considered high. Average posterior probabilities, though useful in themselves when assessing classification accuracy, were also summarized in an entropy statistic that can range from 0 to 1. Higher entropy (closer to 1) indicates better classification accuracy in the dataset and often reflects greater class separation.

3. Removal of Influential Cases. It is possible for extreme response patterns in the data set to disproportionately influence the results of mixture models (e.g., selection of K or size and direction of parameter estimates). Mixture modeling has in fact been noted for its usefulness in outlier identification because sometimes outliers will cause a very small class to be extracted exclusively for the outliers, which may result in estimation problems (Aitkin & Wilson, 1980). Alternately, extreme cases may cause classes to be extracted that do not make substantive sense (e.g., have predicted trajectories that exceed realistic test scores; Jorgensen, 1990).

In fitting the unconditional LCGM model based on students' scores over time on the math sections of the TABE exams, three students' score patterns were found to have the latter effect (extracting nonsensical classes). By examining the original data set, these influential cases were identified as students who entered the program at very high levels (T_0 scores ranged from 10.15 to 12.9) and had only one subsequent score recorded, at T_1 or T_2 , before exiting the program. The outliers were resulting in the existence of a class defined by a trajectory that included predicted TABE scores in excess of the actual possible range (0 to 12.9). In order to preserve interpretability of results for the rest of the sample, the influential cases were removed from the data set and the model was re-run.

4. Conditional LCGM model. Once the best-fitting unconditional LCGM model was determined using the model selection strategies described above, the complexity of the model was increased by adding covariates to the model to predict class membership. The covariates were the five scores X_1 - X_5 previously discussed, each of which represented a different demographic factor being tested for its ability to explain who follows which class-specific learning trajectories. This conditional model, with covariates predicting class membership and class membership in turn determining growth parameters that define an individual's predicted test scores at times T_0 through T_5 , is illustrated by the following path diagram:



The multinomial logistic regression equation for predicting class membership now includes multinomial slopes for person-specific predictors X_{i1} - X_{i5} :

$$p(c_i = k | X_{i1} - X_{i5}) = \frac{\exp(\omega^{(k)} + \beta_1^{(k)} X_{i1} + \beta_2^{(k)} X_{i2} + \beta_3^{(k)} X_{i3} + \beta_4^{(k)} X_{i4} + \beta_5^{(k)} X_{i5})}{\sum_{k=1}^K \exp(\omega^{(k)} + \beta_1^{(k)} X_{i1} + \beta_2^{(k)} X_{i2} + \beta_3^{(k)} X_{i3} + \beta_4^{(k)} X_{i4} + \beta_5^{(k)} X_{i5})}$$

where $\omega^{(K)} = \beta_1^{(K)}, \dots, \beta_5^{(K)} = 0$

In the above equation, $p(c_i = k | X_{i1} - X_{i5})$ refers to the probability of person i belonging to class k , given particular scores on the predictor variables. $\omega^{(k)}$, as in the unconditional model, is a class-specific multinomial intercept, and $\beta_1^{(k)} \dots \beta_5^{(k)}$ refer to class-specific and predictor-specific multinomial slopes, such that they represent the covariates' effects on predicting membership in each class. In order to avoid redundancy, $K-1$ multinomial slopes are estimated for each predictor and the multinomial coefficients for the K th class are fixed at 0.

When the conditional LCGM model was fit to the data, estimated parameters included not only the $K*d$ class-specific growth parameters, T residual variances, and $K-1$ multinomial intercepts mentioned previously, but also $K-1$ multinomial slopes for each covariate predicting class membership (here, for X_1 these would be $\beta_1^{(k)} - \beta_1^{(K-1)}$). These slopes were not on an intuitive metric (they were on a log odds metric), so the values were exponentiated (e.g., $\exp(\beta_1^{(k)})$) and thereby converted to odds ratios. An odds ratio is defined as a ratio of two odds, each of which in turn is a ratio of two probabilities. The odds ratio for a given covariate can be understood as the ratio of the odds of being in the k th class versus a reference (K th class) before and after the covariate value is increased by one. In the case of binary covariates (e.g., unemployed ("0") versus employed ("1")), this would mean the ratio of the odds of being in the k th class vs. the K th class when the covariate value is "0" versus when the covariate value is "1." In the case of continuous covariates (e.g., last grade completed in school), this would mean the ratio of the odds of being in the k th class vs. the K th class at a given level of the covariate versus

when the covariate value increases by one unit.

Given that theory did not suggest one particular reference class for GED student achievement trajectories to which all others could be compared, all odds ratios were re-calculated using each class in turn as a reference class in order to examine how covariates might distinguish membership probability between any two given classes. Odds ratios were determined to be non-significant if their 0.95 confidence intervals overlapped “1” (signifying equal chances of class membership regardless of the covariate value). Odds ratios significant at the $\alpha=0.05$ level were noted and further investigated. Odds ratios that were “marginally” significant ($\alpha=0.10$) were also explored.

5. Associations across processes. The investigation of numbers and orders of class trajectories in unconditional models as well as the investigation of covariate effects in conditional models were conducted separately for math and reading, such that each student was assigned to one class for math and a separate class for reading. To gain a better sense of overall achievement and relationships among growth processes, however, the relationship between these two sets of classes was examined. Doing so addressed the question of whether one’s predicted math growth trajectory affects one’s predicted reading growth trajectory, and vice versa. This analysis was done by fitting a parallel process LCGM model (Nagin & Tremblay, 2001) in *Mplus*. LCGMs for reading and math were both estimated and conditional probabilities of membership in each reading class were estimated given membership in a certain math class, and vice versa.

6. Distal outcomes. After examining patterns and trends that occur while students are enrolled in the GED program, this study examined the impact of such trends on the ultimate success measure of concern: whether students earn their GED certificates. Earning the GED certificate, since it is an event that occurs after the measures included in the TABE score dataset,

was considered a “distal outcome” (Muthén, 2004). Whereas covariates X_1 - X_5 are predictive of class membership, the distal outcome is predicted *by* class membership. The distal outcome in this case, earning versus not earning a GED certificate, was treated as a binary measure. The twelve students in the sample who are still enrolled at MEP, and who therefore have neither “failed” nor “succeeded” in earning their GED certificates, were treated as missing data on the distal outcome measure. Distal outcome data was collected in June 2013, one year after the last TABE test administration included in the study.

Distal outcome analyses were conducted separately for math and reading outcomes. A traditional chi-square analysis was chosen over a recently popularized multi-step approach that involves adding the distal outcome as an auxiliary variable into the LCGM model (Asparouhov, & Muthén, 2013). The latter approach requires that there be variation in the distal outcome within each class (e.g., not every student in a given class attained the GED), and was inappropriate given the relatively low sample size of this study and thus the possibility that there was not variation on the distal outcome within each class. In this study, a contingency table was constructed that defined frequencies of students modally assigned to each class who earned their GED certificates. These frequencies were subjected to a chi-square test of the null hypothesis that a student in one trajectory class was not significantly more likely than a student in another class to pass the GED exam.

Qualitative data

In addition to student test scores and demographic information, qualitative data was provided by MEP in the form of anonymous student remarks about their family backgrounds, educational histories, motivations for entering MEP, experiences in the program, and future goals. This qualitative data provided contextual detail and aided in the interpretation of quantitative results.

Results

Unconditional LCGM: Number and Form of Classes

1. Number of classes. The best fitting number of classes for math and for reading was determined by selecting the number of classes with the lowest Bayesian Information Criteria (BIC) index, allowing for quadratic trajectories in each class. As shown in Table 1 below, $K=3$ classes yielded the lowest BIC for both the math and reading LCGMs. This finding does not imply that there are three fully distinct population subgroups of students within the MEP community; rather, it implies that the individual heterogeneity in learning patterns among MEP students (whether continuously or discretely distributed) is sufficiently accommodated by splitting the students into three groups for analysis. Note that the selected $K=3$ models were not the most complex models tested. The BIC weighs fit against parsimony, so the higher BIC for the four- and five-class models suggests that additional classes beyond three did not meaningfully improve fit when weighed against the additional parameters required.

Table 1
BIC Model Fit Results for Math and Reading, Alternate Numbers of Quadratic Trajectory Classes

Model	Estimated Parameters	BIC (Math)	BIC (Reading)
2-Class Quad	13	1031.205	1317.303
3-Class Quad	17	1004.258	1281.025
4-Class Quad	21	1004.304	1287.700
5-Class Quad	25	1004.694	1286.631

Note. BIC=Bayesian Information Criteria. Quad=quadratic functional form

2. Form of classes. After selecting the number of classes, BIC was used to determine whether a lower-order, more parsimonious functional form than quadratic would improve fit, holding the number of classes constant at three. Linear was determined to be the best fitting functional form for math and for reading, as it resulted in the lowest BIC for each. These results are shown below in Table 2.

Table 2

BIC Model Fit Results for Math and Reading, Alternate Functional Forms for K=3 Classes

Model	Estimated Parameters	BIC (Math)	BIC (Reading)
3-Class Linear	14	992.515	1269.118
3-Class Quad	17	1004.258	1281.025

Note. BIC=Bayesian Information Criteria. Quad=quadratic functional form

Given the model selection results, each class was conceptualized as following a linear trajectory. This means that each class of students was projected to follow a particular academic growth pattern defined by a class-specific intercept (starting level) and a class-specific linear slope (constant rate of growth from one time point to the next, where time points are two months apart on average). The predicted growth trajectories' linear form does not imply that students at MEP learn in a strictly linear fashion; rather, they indicate that a linear form is more appropriate in describing academic growth than a quadratic form.

Conditional LCGM: Class Description and Assignment

1. Class description. After using unconditional LCGMs to determine the number and form of classes, the five covariates were added to both models to predict class membership. Parameter estimates did not vary meaningfully between the unconditional and conditional models, as is to be expected when predictors only affect class membership (Bandein-Roche et al., 1997; Lubke & Muthén, 2007). Therefore, only conditional model parameters estimates are presented here. The class-specific growth parameters estimates are shown in Table 3 for math and Table 4 for reading, and associated graphical representations of the class trajectories are shown in Figures 1 and 2. All parameters except for the linear slope for one reading class were significantly different from zero ($p < .05$). In Figures 1 and 2 below, the observed learning trajectories exhibited by individual students are shown in dashed lines and the class-specific model-implied trajectories, used to parsimoniously summarize this individual variability, are shown in solid lines.

Math

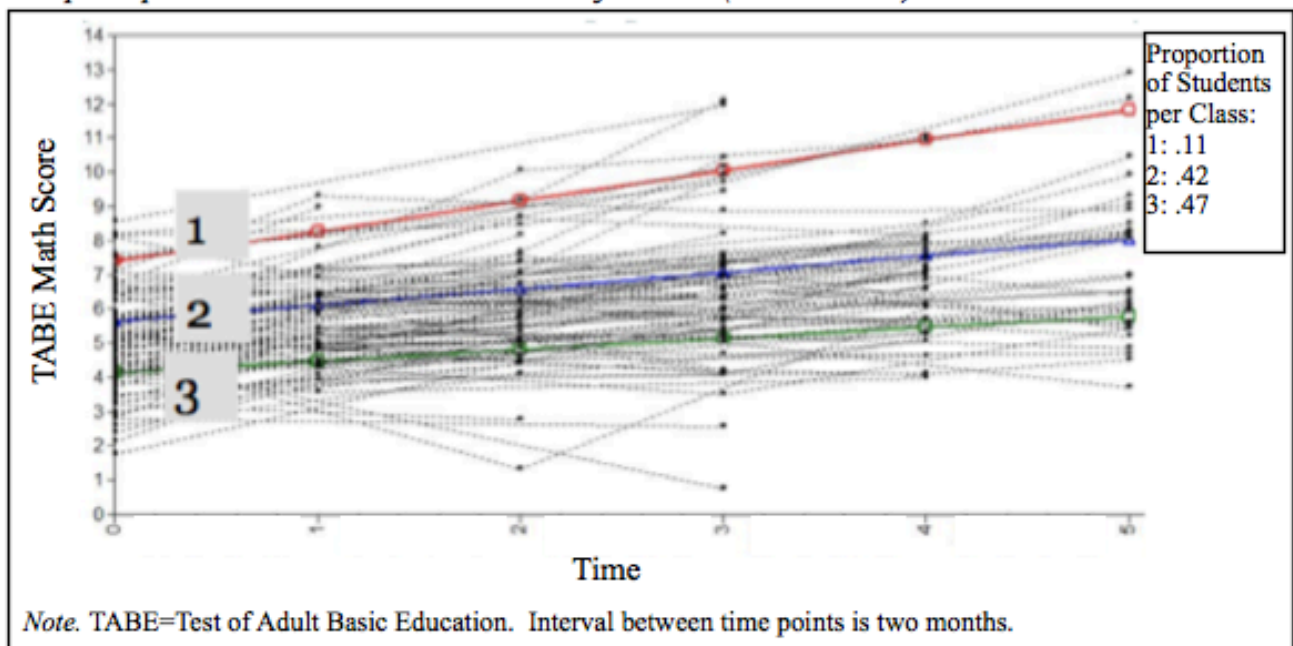
Table 3
Model-Implied Class-Specific Parameters: Math

Class	Intercept (S.E.)	Slope (S.E.)	Proportion of Students
1	7.37* (.30)	.89* (.13)	0.11
2	5.63* (.28)	.48* (.06)	0.42
3	4.15* (.30)	.33* (.07)	0.47

Note. * denotes model parameters significantly different from zero ($p < .05$). Standard errors of parameter estimates are indicated in parentheses.

Figure 1.

Model-Implied Class-Specific Trajectories (Solid Lines) for the 3-Class Math LCGM, Superimposed on Observed Individual Trajectories (Dotted Lines)



As can be seen in Figure 1, the predicted trajectories of the three math classes (solid lines) did not overlap; that is, the students with the highest math level at entry were predicted to remain at the highest level throughout the first 10 months in the program (from T_0 to T_5). Furthermore, the pattern of class-specific intercept and slope pairings reflect the existence of a positive correlation between individuals' intercepts and slopes in the aggregate, such that the students who entered at the highest math level were predicted to grow at the fastest rate.

Class-specific intercepts in Table 3 represent the predicted math level at program entry for students belonging to a particular class. Since TABE test scores are recorded as grade-referenced norms ranging from 0 to 12.9 where the whole number represents the grade in school and the decimal represents the number of months into that grade (such that 12.9=9 months into twelfth grade), intercepts can be interpreted as curricular mastery levels at specific points in students' academic careers. Thus, students assigned to Class 1 enter MEP on average with the math abilities attributed to students mid-way through the third month of seventh grade; students assigned to Class 2, the sixth month of fifth grade; and students assigned to Class 3, the first month of fourth grade.

Linear slopes in Table 3 represent the constant amount of increase on TABE math test scores from one time point to the next, where time points are distanced at two-month intervals. Slopes can be interpreted relative to the growth expected for students following a traditional school curriculum. According to the norm-referenced TABE scores, a traditional student would be expected to increase by 0.9 per nine-month school year (e.g. from 3.0 to 3.9 across the third grade). The slope predicted for students assigned to Class 1, therefore, can be interpreted as meaning that students assigned to Class 1 are predicted to grow academically at a rate of approximately 8.9 months of typical school-based learning over every two-month period, or approximately 4.5 curricular months over every month at MEP. Thus, these students are predicted to accumulate math ability at a rate over quadruple the growth expected in the traditional student population. Students assigned to Class 2 are predicted to grow at a rate of approximately 2.4 curricular months per month; students assigned to Class 3, approximately 1.6 curricular months per month.

In order to ease interpretation, numeric math class labels are supplemented in ensuing analyses by level labels of “Low” (Class 3; 47% of students), “Intermediate” (Class 2; 42%), and “High” (Class 1; 11%). These labels reflect the relative intercepts and slopes of the three classes.

Reading

Table 4

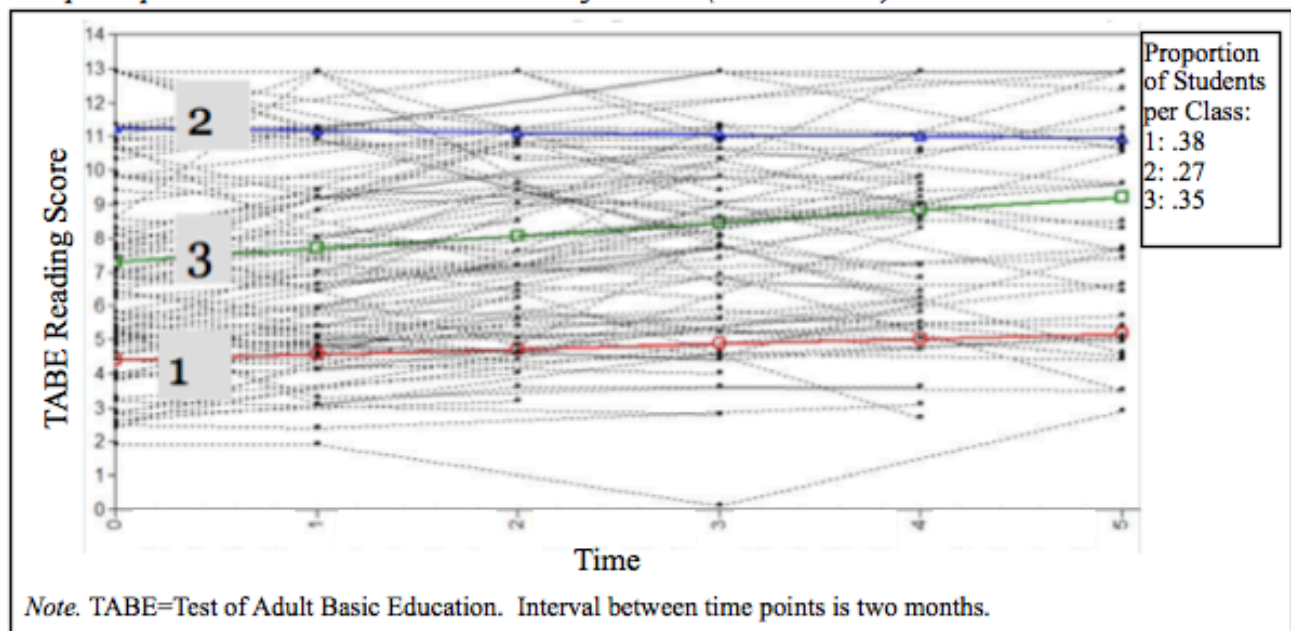
Model-Implied Class-Specific Parameters: Reading

Class	Intercept (S.E.)	Slope (S.E.)	Proportion of Students
1	4.40* (.24)	.15* (.05)	0.38
2	11.25* (.32)	-.063 (.15)	0.27
3	7.30* (.44)	.38* (.13)	0.35

Note. * denotes model parameters significantly different from zero ($p < .05$). Standard errors of parameter estimates are indicated in parentheses.

Figure 2.

Model-Implied Class-Specific Trajectories (Solid Lines) for the 3-Class Reading LCGM, Superimposed on Observed Individual Trajectories (Dotted Lines)



As with math, class-specific predicted reading trajectories (solid lines) shown in Figure 2 did not overlap, suggesting that the students who enter MEP at the highest reading levels are still reading at a higher level than their peers after 10 months in the program. Unlike math, however, higher intercepts were not accompanied by more positive linear slopes for all classes. While the

students entering at approximately a seventh grade reading level (Class 3) were predicted to advance at a faster rate than those entering around a fourth grade reading level (Class 1), the students in Class 2 whose entry reading levels were, on average, equal to students in the second month of eleventh grade had a non-significant growth rate ($p > .05$). This is likely a ceiling effect due to the fact that the TABE test scores are capped at 12.9, so the students reading at the highest levels did not have “room to grow” in reading, but rather remained at MEP to improve upon other subjects such as math.

As with math, numeric reading class labels are supplemented in ensuing analyses by level labels in order to ease interpretation. Since reading class slopes varied substantially across classes, labels include slope descriptors. Reading classes are henceforth summarized as “Low-Slowly Increasing” (Class 1; 38% of students), “High-Stable” (Class 2; 27%), and “Intermediate-Increasing” (Class 3; 35%).

2. Assessment of classification accuracy. After estimating the discrete class trajectories in math and in reading, each individual was assigned to her most likely class in each subject. This assignment was based on posterior probabilities, as previously discussed, which are defined as the probability of an individual being in each class based on her pattern of scores across time and the model estimates. Each student was assigned to the class for which she had the highest posterior probability of membership.

Classification accuracy, as previously discussed, was calculated based on average posterior probabilities of belonging in each class, and is summarized in Tables 5 and 6 below for math and reading, respectively. Rows refer to the classes to which individuals *were* modally assigned, and columns refer to the classes to which those individuals *could* theoretically have been assigned. Thus, the probability in row 1, column 2 of Table 5 (.03) represents the average probability that individuals assigned to Class 1 in math had of belonging to Class 2. Diagonals

in Tables 5 and 6 represent individuals' average posterior probability of belonging to the classes to which they were assigned. Probabilities in the off-diagonals represent the average posterior probabilities of belonging to classes to which they were not assigned. The high diagonal probabilities and low off-diagonal probabilities in Tables 5 and 6 evidence a quite high degree of classification accuracy in the LCGMs for math and reading; that is, given a student's observed series of TABE scores she could be classified into one of the three model-implied classes on math and on reading with little uncertainty.

Table 5

Posterior Probabilities of Class Assignment, by Modally Assigned Class: Math

Modally Assigned Class	Class 1 (High)	Class 2 (Intermediate)	Class 3 (Low)
1	.97	.03	.00
2	.02	.89	.09
3	.00	.08	.92

Note. Entropy=.80

Table 6

Posterior Probabilities of Class Assignment, by Modally Assigned Class: Reading

Modally Assigned Class	Class 1 (Low-Slowly Increasing)	Class 2 (High-Stable)	Class 3 (Intermediate-Increasing)
1	.97	.00	.03
2	.00	.97	.03
3	.05	.03	.92

Note. Entropy=.88

As a single-number summary of these classification accuracy tables, the entropy statistics for math and reading were .80 and .88, respectively, in the conditional LCGMs. This reflected a quite high degree of classification accuracy and class separation for model-implied trajectories in the models.

Conditional LCGM: Prediction of Class Membership

The ability of each of the five covariates to predict class membership was assessed by calculating odds ratios for each covariate across each pair of math classes and, separately, each pair of reading classes. Odds ratios, as previously discussed, can be understood as the ratio of

the odds of being in the k th class versus a reference K th class before and after the covariate value is increased by one unit (for continuous covariates) or changed from 0 to 1 (for binary covariates). Tables 7 thru 10 below show odds ratio results for each covariate in math and in reading, across each pair of reference classes K and comparison classes k . An odds ratio is denoted as significant if its 95% confidence interval did not include 1 and marginally significant if its 90% confidence interval did not include 1 (here, 1 represents no change in the odds of membership in class k versus class K).

Math

Table 7

Conditional LCGM for Math: Odds Ratios versus Class 3 (Low)

Comparison Class	X ₁ Last Grade	X ₂ Age at Entry	X ₃ Language	X ₄ Employment	X ₅ Head of Household
2 (Intermediate)	.79	.95	.44	3.83	.49
1 (High)	1.19	.92*	2.22	9.41*	1.30

Note. * indicates significance ($p < .05$). † indicates marginal significance ($p < .1$).

Table 8

Conditional LCGM for Math: Odds Ratios versus Class 2 (Intermediate)

Comparison Class	X ₁ Last Grade	X ₂ Age at Entry	X ₃ Language	X ₄ Employment	X ₅ Head of Household
1 (High)	1.50	.95	5.04†	2.46	2.64

Note. * indicates significance ($p < .05$). † indicates marginal significance ($p < .1$).

Three covariates were significant in predicting students' model-implied math class membership: age at entry, employment status, and language spoken at home.

As shown in Table 7, age at entry yielded a significant odds ratio of .92 for the High class compared to the Low class, suggesting that for each one-year increase in a student's age, her likelihood of belonging to the High vs Low math class *decreased* by 8 percent. The significance of this covariate did not hold for the Medium vs Low or for High vs Medium (Table 8) class comparisons, suggesting that while age may be valuable in distinguishing between the highest-

and lowest-achieving math students, it is not necessarily predictive of more subtle differences in math ability.

Employment status, like age at entry, yielded a significant odds ratio for the High vs Low math class comparison, but not for more subtle distinctions between student math abilities. The odds ratio of 9.41 indicates that students who held jobs while at MEP were over 9 times more likely than students who were unemployed of being assigned to the High class vs the Low class. This suggests that employment is positively predictive of a students' incoming math ability as well as her subsequent learning growth in math.

Language yielded a significant odds ratio of 5.038 for the High vs Low class comparison in Table 7. This means that a woman who does *not* speak English at home ($X_3=1$) was 5 times *more likely* to be assigned to the High vs Medium math class than is a woman who does speak English at home.

Reading

Table 9

Conditional LCGM for Reading: Odds Ratios versus Class 1 (Low-Slowly Increasing)

Comparison Class	X ₁ Last Grade	X ₂ Age at Entry	X ₃ Language	X ₄ Employment	X ₅ Head of Household
3 (Intermediate-Increasing)	.86	.96	.12*	.77	.47
2 (High-Stable)	.97	.94	.17†	2.27	1.87

Note. * indicates significance ($p < .05$). † indicates marginal significance ($p < .1$).

Table 10

Conditional LCGM for Reading: Odds Ratios versus Class 3 (Intermediate-Increasing)

Comparison Class	X ₁ Last Grade	X ₂ Age at Entry	X ₃ Language	X ₄ Employment	X ₅ Head of Household
2 (High-Stable)	1.12	.98	1.42	2.93	3.96*

Note. * indicates significance ($p < .05$). † indicates marginal significance ($p < .1$).

Two covariates were significant in predicting students' model-implied reading class membership: language spoken at home and head of household status.

Language spoken at home yielded odds ratios significantly lower than 1 for the Intermediate-Increasing vs Low-Slowly Increasing reading class comparison as well as the High-Stable vs Low-Slowly Increasing reading class comparison, as shown in Table 9. This indicates, sensibly, that students who did not speak English at home had a much *lower* probability of belonging to the model-implied Intermediate-Increasing or High-Stable classes versus the Low-Slowly Increasing class than did their English-speaking peers. Specifically, the odds that a student who did not speak English at home would be assigned to the Intermediate-Increasing or High-Stable reading class rather than to the Low-Slowly Increasing reading class were about 88 percent and 83 percent lower, respectively, than were these odds for a peer who did speak English at home.

Head of household status had the opposite effect on predicted reading class membership; that is, being a female single head of household (FSHH) positively predicted assignment to the highest reading class. As shown in Table 10, an FSHH woman's odds of being assigned to the High-Stable vs Intermediate-Increasing reading class were nearly 4 times higher than the same odds for a woman who was not an FSHH.

Associations Across Processes

The parallel process LCGM was constructed by combining the best-fitting math and reading LCGMs and regressing class membership for each process on the other. From this model, joint probabilities as well as conditional probabilities of membership were calculated.

1. Joint probabilities of math and reading class membership. Table 11 shows students' probability of belonging to each combination of math and reading classes.

Table 11
Joint Probabilities of Math and Reading Class Membership

	Math Class 3 (Low)	Math Class 2 (Intermediate)	Math Class 1 (High)
Reading Class 1 (Low-Slowly Increasing)	.32	.04	.00
Reading Class 3 (Intermediate-Increasing)	.07	.27	.02
Reading Class 2 (High-Stable)	.03	.15	.09

1. Conditional probability of reading class given math class. Table 12 shows an individual's probability of belonging to a particular reading class (columns) given her membership in a particular math class (rows). As can be observed in the table, given an individual's assignment to a particular math class, she is highly likely to be assigned to a class of the same relative standing on reading. A student assigned to the High math class, for example, has a .83 probability of being assigned to the High-Stable reading class. There is slightly more uncertainty in whether women from the Medium math class will be assigned to the Intermediate-Increasing (.59) or High-Stable (.33) reading class. It is important to note, however, that membership in the High-Stable reading class (27 women) far exceeded membership in the High math class (11 women), thus precluding full alignment of relative class standings across subjects.

Table 12
Conditional Probabilities of Reading Class Membership, Given Math Class

Math Class	Class 1 (Low-Slowly Increasing)	Class 3 (Intermediate-Increasing)	Class 2 (High-Stable)
Class 3 (Low)	.76	.16	.08
Class 2 (Intermediate)	.08	.59	.33
Class 1 (High)	.00	.17	.83

2. Conditional probability of math class given reading class. Table 13 shows an individual's probability of belonging to a particular math class (columns) given her membership in a particular reading class (rows). Membership in the Low-Slowly Increasing or Intermediate-

Increasing reading class is clearly associated with membership in a math class at the same relative standing (probabilities of 0.89 and 0.76, respectively). This relationship, however, is not apparent for those in the High-Stable reading class. When interpreting this disparity, it is important to again consider the relative sizes of the classes: while the model-implied High-Stable reading class had 27 members, the model-implied High math class had only 10 members. Thus, the probability of a member of the highest reading class being assigned to the highest math class was capped at $10/27$, or 0.37. The observed probability of 0.33 is thus close to its maximum value. Still, when compared to the .83 probability of High-Stable reading class membership given High math class membership, it is apparent that there is much more variability in math performance among strong readers than vice versa. If all that is known about a woman is that she is high-performing in math, it is quite certain that she will be high-performing in reading; however, if all that is known is that she is high-performing in reading, her math abilities are less easily inferred.

Table 13

Conditional Probabilities of Math Class Membership, Given Reading Class

Reading Class	Class 3 (Low)	Class 2 (Intermediate)	Class 1 (High)
Class 1 (Low-Slowly Increasing)	.89	.11	.00
Class 3 (Intermediate- Increasing)	.19	.76	.05
Class 2 (High- Stable)	.12	.55	.33

Distal Outcomes

Tables 14 and 15 show the number of students assigned to each math and reading class, respectively, who by June 2013 had dropped out of MEP, passed the GED, or were still enrolled in the program. Distal outcomes for math and for reading were each subjected to a chi-square test of independence. “Still enrolled” students were treated as missing data, such that the test

focused solely on students in each class who had passed the GED versus dropped out. The chi-square test was thus a 3x2 design (3 classes, 2 GED status outcomes). The chi-square tests yielded results of 20.092 ($p < .001$) and 26.732 ($p < .001$) for math and reading, respectively. The null hypotheses that math class membership and GED status are independent and that reading class membership and GED status are independent were thus both rejected.

Table 14

GED Success Distal Outcome Predicted by Math Class Membership: Frequencies and Proportions

Distal Outcome	Class 3 (Low)	Class 2 (Intermediate)	Class 1 (High)
Dropped Out	24 (.53)	9 (.22)	0 (.00)
Passed GED	13 (.29)	28 (.68)	10 (1.0)
Still Enrolled	8 (.18)	4 (.10)	0 (.00)

Note. Distal outcome data was collected from MEP records in June 2013. Proportions are shown in parentheses.

Table 15

GED Success Distal Outcome Predicted by Reading Class Membership: Frequencies and Proportions

Distal Outcome	Class 1 (Low-Slowly Increasing)	Class 3 (Intermediate-Increasing)	Class 2 (High-Stable)
Dropped Out	21 (.57)	11 (.30)	1 (.04)
Passed GED	9 (.24)	20 (.56)	26 (.96)
Still Enrolled	7 (.19)	5 (.14)	0 (.00)

Note. Distal outcome data was collected from MEP records in June 2013. Proportions are shown in parentheses.

In sum, the chi-square analysis of distal outcomes with regard to students' assigned classes on math and on reading resulted in strong evidence of systematic differences between classes in terms of future success. In other words, students' entry levels and subsequent learning trajectories within their first ten months at MEP were strongly associated with their likelihood of ultimately passing the GED versus dropping out by June 2013, when the distal outcome data was collected.

Discussion

This thesis was guided by four primary research questions: description of heterogeneity in math and reading academic growth; prediction of student learning by pre-existing demographic factors; associations across math and reading growth trajectories; and associations between growth trajectories and GED attainment. The results pertaining to each of these components are discussed below.

Research Question 1: Math and Reading Trajectory Class Description

A key objective of this thesis was to describe student learning at MEP in a way that accommodated the vast heterogeneity among students' entry levels and rates of growth. Model selection results suggest that a parsimonious representation of this individual heterogeneity is accomplished with three distinct LCGM classes in math and in reading, each with unique intercepts and rates of change.

In math, classes with more positive slopes had higher intercepts. Greater math ability at program entry seems to facilitate faster learning, rather than the alternate possibility that it contributes to student complacency.

In reading, this facilitation effect—higher entry levels resulting in faster academic growth—held true between the Low-Slowly Increasing and Intermediate-Increasing classes, but not between the Intermediate-Increasing and High-Stable classes. This is likely due, however, to the fact that the average incoming reading level of students assigned to the highest class was very close to the maximum possible TABE score. MEP students, it seems, may often enter with reading skills that outpace their skills in other subjects such as math, and remain in the program for the enhancement of these subjects rather than reading.

Trajectory classes extracted in this thesis need not be considered literal subpopulations; it is not necessary to conceptualize “true” classes of learners within any student population.

Rather, the use of LCGM to divide students into groups serves to parsimoniously summarize prominent features of the heterogeneity in growth among the student population, even if this heterogeneity is in fact continuously distributed (e.g., Sterba, Baldasaro, & Bauer, 2012). This thesis thus demonstrates the use of LCGM techniques to clarify the nature of heterogeneity in academic settings, even if that heterogeneity is not in the form of “true” classes in a literal sense. Furthermore, the analysis of predictors and distal outcomes related to these model-implied divisions allows for a more nuanced understanding of the factors and processes underlying achievement than is incorporated in traditional whole-population longitudinal analyses (Sterba & Bauer, in press).

On a whole, the model-implied linear slopes suggest that students at MEP learn at faster rates than traditional student populations, particularly in math. In terms of norm-referenced K-12 student growth, the model-implied linear slopes suggest that MEP students in the Low, Intermediate, and High math classes grow at a rate of 1.6, 2.4, and 4.5 curricular months, respectively, for every month they spend at MEP. There are several possible explanations for the relative success of MEP compared to traditional classrooms. First, students at MEP are much older than traditional students learning at the K-12 level, and might therefore be more diligent in their learning and/or experienced in integrating new knowledge. Second, MEP emphasizes tailored instruction with substantial individual tutoring time, which may facilitate faster growth than learning environments in which nearly all instruction is class-wide. Third, some self-selection effect may exist among the MEP student population, as these students—unlike traditional elementary and middle schoolers—have voluntarily opted to attend and might therefore be more motivated to maximize their time and more invested in their learning.

In reading, the model-implied slope of the Intermediate-Increasing class also suggests that students learn more quickly than anticipated by traditional curriculums (1.9 curricular

months for every month at MEP). The average growth among students in the Low-Slowly Increasing and High-Stable classes, however, fell short of anticipated growth in traditional classrooms. The quick growth of the Intermediate-Increasing class might be explained by the possibilities listed above. The non-significant linear slope of the High-Stable class is likely due to a ceiling effect; these students entered MEP, on average, at a reading level of 11.25 (2.5 months into eleventh grade) where the maximum level is 12.9. The sub-standard growth among the Low-Slowly Increasing class is somewhat more difficult to explain, especially in light of the fact that these students still exhibited accelerated growth in math. Possible explanations include general unfamiliarity with the English language (many students are non-native English speakers) and underlying learning disabilities such as dyslexia that inhibit reading more harshly than math.

Research Question 2: Prediction of Trajectory Class Membership

Odds ratio results in the conditional LCGM model suggest some prediction effects that are consistent with prior theory as well as some that differ from prior theory. The latter may be due in part to the tendency of prior research to define predictor variables in broad terms such as “situational barriers” or “self-efficacy” (versus this thesis, which used narrow demographic factors), and also in part to the particular population of students at MEP.

The non-significance of last grade completed in predicting students’ class assignment counters Cervero’s (1983) previous finding that students with more schooling were more likely to succeed in GED preparation programs and pass the exam. It is possible that the non-significance of prior schooling in this study was due to the demographics of MEP’s particular student population: most students are living below poverty line and attended schools in under-resourced districts. In these districts, school attendance might not align reliably with curricular mastery and therefore would not systematically predict students’ TABE scores upon entering MEP. Since class membership assignments are dependent upon both entry level (intercept) and

learning speed (linear slope), it is also important to note that the non-significance of last grade completed indicates that students' prior school experience did not seem to "prime them" for learning at MEP. Thus, just as the years they spent in formal school environments were not associated with material mastery, they also were not associated with an ability to absorb new information quickly.

The significance of age at program entry in predicting math class membership, however, suggests some potential impact of school on GED program achievement: the longer a student has been out of school, the more her math skills have declined. Alternately, it is possible that the effect of age is not mediated by information retention (or lack thereof), but rather is due to a general decline in brain plasticity and learning speed as a student ages. The latter alternative seems more in line with prior studies by Herring (2012) and King (2002) that found older adults to be less likely to enroll and persist in GED programs.

While Herring (2012) named childcare as a "situational barrier" that might undermine success in GED preparation programs and King (2002) cited family obligations as an obstacle to GED attainment, a proxy for these variables used in this thesis—female single head of household (FSHH) status—was related to higher reading achievement and had no bearing on math achievement. MEP's provision of childcare for its students may account for this difference. Furthermore, it is possible that FSHH status requires women to acquire and maintain higher reading skills in order to communicate with their children's schools, medical providers, etc.

Employment status was similarly anticipated, based on prior research, to negatively affect students' academic growth due to time conflicts. The results of this analysis, however, suggest otherwise: employment contributed positively to a student's likelihood of belonging to the highest-achieving math class, and did not affect her reading class assignment. It appears that women consider and resolve time conflicts prior to making the commitment to attend MEP, and

that employment—via job requirements or experience with paychecks and taxes—may actually aid students’ understanding of math and numbers. A reverse effect, however, might also account for this relationship: students with good math skills may be more likely to gain employment, rather than employment preceding the development of better math skills.

Speaking a language other than English at home, while hypothesized to be associated with lower initial ability and slower improvement in both math and reading, showed mixed effects in this analysis. While it was indeed associated with membership in the lowest reading class, it was positively associated with membership in the highest math class. Several possibilities exist for this effect, though deeper exploration is needed to externally validate any. First, difficulty experienced in reading and other verbally intensive subjects might cause non-native English speakers to overemphasize mathematical learning and devote more time to mastery of nonverbal skills. Second, non-native English speaking students might have grown up in homes and cultures that strongly promoted STEM disciplines and thus endowed them with mathematical foundations outside of school. These students might then attend MEP solely for assistance with reading and not math. Third, the effect of language may differ across levels of another predictor variable in the dataset; for example, non-native English speakers may have overwhelmingly been employed or younger, factors that were related to assignment in higher math classes, and the effect of language on class membership may not have indeed been additive beyond age, employment, etc. In this thesis, conditional LCGMs included main effects of the five predictors but did not include interactions among predictors; such interaction effects could be a valuable area for future research investigation.

Research Question 3: Associations Across Math and Reading Trajectory Classes

The conditional probabilities of membership in a certain math class given membership in a certain reading class, and vice versa, suggest strong relationships between subjects in terms of

student achievement. This is interesting in light of the non-significance of prior schooling in determining students' class assignments. It seems that GED students' initial performance and progression academically is somewhat generalized across subjects, albeit not determined by the extent of their formal schooling. As previously mentioned, the non-significance of formal schooling is unlikely to hold true across all student populations, but may indicate failures in the low-income school districts from which most MEP students hail. Some general intelligence factor, therefore, may underlie the consistency across subjects of students' relative standing among their peers.

The relative predictive value of membership in the highest math class versus the highest reading class was additionally interesting. While membership in the High math class rendered membership in the High-Stable reading class very likely, the reverse relationship did not hold. This suggests that more variation in math ability exists among high achievers in reading than does variation in reading ability among high achievers in math.

Research Question 4: GED Attainment as a Distal Outcome of Class Membership

Distal outcome analyses suggested a strong effect of both math and reading class membership on students' future outcomes in terms of GED attainment. This might at first appear intuitive: students in the highest-performing classes on math and on reading are the most likely to pass the GED exam. It is important to consider, however, that the analysis only included TABE test scores from students' first ten months at MEP. Students' distal outcome status—passing the GED versus dropping out of MEP—was collected in June 2013, twelve months after the last recorded TABE score. For some students whose score collection began as early as 2006, the distal outcome data represented their status nearly seven years after entering MEP. Thus, it was possible that students in the lowest-performing classes persisted after their first ten months

and earned the GED certificate, albeit not as quickly as their peers who showed higher performance in the first ten months.

The highly significant chi-square results suggest, to the contrary, that students who entered MEP at low academic levels in math and in reading and who subsequently grew at lower rates than their peers largely did *not* persist in the program enough to attain the GED. This provides evidence that GED programs may not provide “level playing fields” for all attendees, and that a student’s ability at entry and growth within the first year are highly predictive of whether she will ultimately succeed. This sheds light upon the implications of Lewis’ (2009) finding that poor math and reading skills upon entry into GED programs adversely affect self-efficacy: the adverse effect may not wear off with subsequent effort and learning, but rather may undermine a student’s likelihood of reaching her GED goals. It may be wise for educators and policymakers to explore additional intervention services aimed at improving self-efficacy among particularly low-performing students entering GED programs in order to help these students persist in the programs and have steeper positive learning trajectories.

Limitations and Future Directions

This thesis used LCGM techniques to describe and predict academic growth at MEP using pre-collected questionnaire and test data. The analyses presented are not aimed at positing definitive causal connections or fixed categorizations of student learning. Stronger causal inferences about the mechanisms of change through which pre-existing demographic variables affect math and reading achievement trajectories would be facilitated through a randomized experiment. In such an experiment, women could be matched on demographic factors and one factor could be manipulated (e.g., half of a sample of women who are not native English speakers could be given intensive English language instruction prior to GED program

enrollment). The women could then be subjected to identical GED training programs, such that the effect of the one variable factor on student learning could be isolated.

Data sparsity posed a potential problem in this thesis, in the form of individually varying measurement intervals (differing lapses of time between test administrations), intermittent missing data (some students lacked TABE scores for multiple time points), and dropout. Though a binning approach was used to address individually varying measurement intervals and the maximum likelihood estimation method used to fit LCGMs accounted for MAR missingness, it is possible that the LCGM model was somewhat compromised by non-randomly missing data involving dropout. The application of LCGM to student populations could be optimized by future longitudinal studies that involve more controlled data collection in consistent intervals across all subjects (rather than ex post facto data attainment).

Conclusion

Conventional wisdom holds that no two students—whether kindergarteners or forty-year-old GED seekers—are the same. The nuances of how, why, and to what extent they differ, however, are less intuitive. Recent advances in statistical methods, including Latent Class Growth Modeling, can help improve educators', policymakers', and other stakeholders' understanding of academic differences.

This thesis moves beyond traditional treatments of GED students as a singular population, and illustrates how conceptualizing complex inter-individual variability in terms of a parsimonious number of latent trajectory classes may help illustrate and explain variations in observed academic achievement. Accurately accommodating student differences has long been a challenge for educators, policymakers, and researchers alike. Latent Class Growth Modeling may be a valuable tool for not only modeling GED students' achievement but also for understanding learner populations in other settings. This thesis has further demonstrated that by

describing student growth in terms of multiple predicted trajectories with unique combinations of entrance level and rate of change, it is possible to identify factors that predict which students will follow which growth patterns and, by extension, which students will achieve their ultimate goals.

MEP students are bound by common goals: “[We’re all] having the same struggles, wanting the things I want.”³ “We all trying to press our way to go forward.” The process of going forward, however, is far from universal. “I’m just pacing myself. Other women are moving faster, [but] I gotta take my time, I can’t speed up.” “Sometimes it make me sad when the same people that came in with me, they done graduated I’m like how’d they do that so quick.”

Some students will persist. Others will not. This thesis helps unravel the complexity of why—and how—students committed to a common goal experience such different levels of progress and success.

³ Student quotations are derived from personal interviews conducted by the author in March 2012. Fourteen women total were interviewed, twelve of whom were also in the sample used for the quantitative analyses presented in this thesis.

References

- Aitkin, M. & Wilson, G. (1980). Mixture models, outliers and the EM algorithm. *Technometrics*, 22, 325-331.
- Asparouhov, T. & Muthén, B.O. (2013). Auxiliary variables in mixture modeling: 3-step approaches using *Mplus*. *Mplus Web Notes No. 15*. Retrieved from www.statmodel.com.
- Bandeen-Roche, K., Miglioretti, D. L., Zeger, S. L., & Rathouz, P. J. (1997). Latent variable regression for multiple discrete outcomes. *Journal of the American Statistical Association*, 92, 1375-1386.
- Bujack, L. K. (2012). *Predicting student persistence in adult basic education using interaction effects among academic self-efficacy and students participation and academic variables*. (Doctoral dissertation, Capella University). Retrieved from <http://search.proquest.com.proxy.library.vanderbilt.edu/docview/1013836709/13E2DC8286463651A0E/9?accountid=14816>
- Cervero, R. M. (1983, April). *A national survey of GED test candidates: Preparation, performance, and 18 month outcomes*. Paper presented at the Annual Meeting of the American Educational Research Association, Montreal, Quebec.
- Crissey, S. R., & Bauman, K. J. (2012). *Measurement of high school equivalency credentials in census bureau surveys* (SEHSD Working Paper No. 2012-3). Retrieved from U.S. Census Bureau Social, Economic, and Housing Statistics Division website: http://www.census.gov/hhes/socdemo/education/data/cps/GED_wp2012-3.pdf
- Davis, C.A., Mottern, R., & Ziegler, M.F. (2010). "It's like a prison without bars": Experiences in a mandatory GED program. In Gandy, P., Tieszen, S., Taylor-Hunt, C., Flowers, D., and Sheared, V. (2010). 51st Adult Education Research Conference (AERC) and 3rd Western

- Region Research Conference on the Education of Adults (WRRCEA) Joint Conference Proceedings, (pp. 115-121). Sacramento, CA: California State University
- Ewert, S. (2012, February 27). GED recipients have lower earnings, are less likely to enter college [Web log post]. Retrieved from <http://blogs.census.gov/2012/02/27/ged-recipients-have-lower-earnings-are-less-likely-to-enter-college/>
- GED Testing Service. (2012). *2012 annual statistical report on the GED test*. Retrieved from <http://www.gedtestingservice.com/educators/historical-testing-data>
- Heckman, J. J., & Rubinstein, Y. (2001). The importance of noncognitive skills: Lessons from the GED testing program. *The American Economic Review*, *91*(2), 145-149. Retrieved from <http://search.proquest.com.proxy.library.vanderbilt.edu/docview/233039004/13F4930543C68F40B10/1?accountid=14816>
- Heckman, J. J., Humphries, J. E., & Mader, N. S. (2010). *The GED*. (National Bureau of Economic Research Working Paper). Retrieved from U.S. National Bureau of Economic Research website: http://www.nber.org/papers/w16064.pdf?new_window=1
- Herring, A. H. (2012). *An examination of motivational factors in high school dropouts participating in general educational development degree programs*. (Doctoral dissertation, Auburn University). Retrieved from <http://etd.auburn.edu/etd/bitstream/handle/10415/3099/FINAL.pdf?sequence=2>
- Jorgensen, M. (1990). Influence-based diagnostics for finite mixture models. *Biometrics*, *46*, 1047-1058.
- King, C. (2002). Barriers affecting GED participation among recent high school dropouts. *Adult Basic Education*, *12*(3), 145-156. Retrieved from <http://search.proquest.com.proxy.library.vanderbilt.edu/docview/227581183/13F4AB8F9264E04A8D6/2?accountid=14816>

- Leininger, L. J., & Kalil, A. (2008). Cognitive and non-cognitive predictors of success in adult education programs: Evidence from experimental data with low-income welfare recipients. *Journal of Policy Analysis and Management*, 27(3), 521-535.
- Lewis, T. M. (2009). *Identifying ways to improve learner persistence in GED programs* (Doctoral dissertation, Capella University). Retrieved from <http://gradworks.umi.com/33/87/3387216.html>
- Lubke, G., & Muthén, B. O. (2007). Performance of factor mixture models as a function of model size, covariate effects, and class-specific parameters. *Structural Equation Modeling*, 14(1), 26-47.
- McGraw Hill. (2013). TABE Overview. Retrieved on June 1, 2013 from <http://www.ctb.com/ctb.com/control/productFamilyViewAction?productFamilyId=608&p=products>
- McLachlan, G. & Peel, D. (2000). *Finite mixture models*. New York: Wiley.
- Mercy Education Project. (2013, Fall). *Mercy Education Project Annual Report 2012-2013*. Detroit, MI: Mercy Education Project.
- Moore, R. T., & Davies, J. A. (1984). Predicting GED scores on the bases of expectancy, valence, intelligence, and pretest skill levels with the disadvantaged. *Educational and Psychological Measurement*, 44(2), 483-490.
- Muthén, B. O. (2004). Latent variable analysis: Growth mixture modeling and related techniques for longitudinal data. In D. Kaplan (ed.), *Handbook of quantitative methodology for the social sciences* (pp. 345-368). Newbury Park, CA: Sage Publications.
- Muthén, L. K. & Muthén, B. O. (1998-2013). *Mplus Users Guide (7th Ed.)*. Los Angeles, CA: Muthén & Muthén.
- Muthén, B. O., & Shedden, K. (1999). Finite mixture modeling with mixture outcomes using the

- EM algorithm. *Biometrics*, 55, 463–469.
- Nagin, D. S. (1999). Analyzing developmental trajectories: A semiparametric, group-based approach. *Psychological methods*, 4(2), 139-157.
- Nagin, D. S. (2005). *Group-based modeling of development*. Cambridge, MA: Harvard Press.
- Nagin, D. S. & Tremblay, R.E. (2001). Analyzing developmental trajectories of distinct but related behaviors: A group-based method. *Psychological Methods*, 6, 18-34.
- Nylund, K. L., Asparouhov, T., & Muthén, B. O. (2007). Deciding on the number of classes in latent class analysis and growth mixture modeling: A monte-carlo simulation study. *Structural Equation Modeling*, 14, 535-569.
- Ritchie, S. J., & Bates, T. C. (2013). Enduring links from childhood mathematics and reading achievement to adult socioeconomic status. *Psychological Science*, 24(7), 1301-1308.
- Ruzzi, B. B., & Kraemer, J. (2006). *Academic programs in alternative education: An overview*. (National Center on Education and the Economy Working Paper). Retrieved from <http://www.ncee.org/wp-content/uploads/2010/04/AcademicProg.pdf>
- Shaffer, D. R. (2008). *Social and personality development* (6th ed.). Belmont, CA: Wadsworth.
- Sterba, S. K. (In press). Fitting nonlinear latent growth models with individually-varying time points. *Structural Equation Modeling*.
- Sterba, S. K., Baldasaro, R. E., & Bauer, D. J. (2012). Factors affecting the adequacy and preferability of semiparametric groups-based approximations of continuous growth trajectories. *Multivariate Behavioral Research*, 40, 590-634.
- Sterba, S. K., & Bauer, D. J. (In press). Predictions of individual change recovered with latent class or random coefficient growth models. *Structural Equation Modeling*.
- Tofghi, D., & Enders, C. K. (2007). Identifying the correct number of classes in growth mixture models. In G. R. Hancock & K. M. Samuelsen (Eds.), *Advances in latent*

variable mixture models (pp. 317-341). Greenwich, CT: Information Age.

Tyler, J. H., & Lofstrom, M. (2009). Finishing high school: Alternative pathways and dropout recovery. *The Future of Children, 19*(1), 77-103.

U.S. Bureau of Labor Statistics. (2012). *Employment status of the civilian noninstitutional population 25 years and over by educational attainment, sex, race, and Hispanic or Latino ethnicity*. Retrieved from <http://www.bls.gov/cps/cpsaat07.pdf>

U.S. Census Bureau. (2011). *Selected social characteristics in the United States, 2007-2011 American Community Survey 5-year estimates*. Retrieved from <http://factfinder2.census.gov>