

Examining the Multiple Trajectories Associated with Dropping Out of High School: A Growth Mixture Model Analysis¹

Alex J. Bowers² & Ryan Sprott

Teachers College, Columbia University & The University of Texas at San Antonio

ABSTRACT

Historically, students who fail to graduate from secondary school are considered as a single category of school dropouts. However, emerging literature indicates that there may be multiple subgroups of high school dropouts, termed dropout typologies. The purpose of this study was to assess the extent to which different typologies of dropouts were present in a large national dataset, and to estimate the influence of the known covariates of dropping out on each of the subgroups. A growth mixture model was estimated using the ELS:2002 dataset and non-cumulative GPA during the first three semesters of high school. The model identified two main subgroups associated with dropping out which accounted for 24.6% of the sample but contained 91.8% of the dropouts.

Keywords: Dropouts, dropout attitudes, dropout characteristics, dropout identification, dropout prediction, dropout prevention, dropout problems, dropout research, at risk students, high school, secondary education, high school graduates, ELS:2002, national surveys, longitudinal studies, growth mixture modeling

INTRODUCTION

Purpose & Background:

Students' dropping out of high schools in the United States is a well-known and pervasive problem. In comparison to students who graduate, students who do not complete high school have lower overall expected life outcomes, including lower lifetime earnings, lower rates of employment, decreased health, and higher incarceration rates (Moretti, 2007; Muenning, 2007; Rouse, 2007; Waldfogel, Garfinkel, & Kelly, 2007). Nationwide, estimates of the percentage of students who fail to complete high school range from nearly 20% to higher than 50% for many large urban high schools (Cataldi, Laird, & KewalRamani, 2009; Swanson, 2004). However, for school leaders, how does one identify the students most likely to drop out so as to provide those students additional resources such as mentoring, tutoring or school studying and behavior interventions (Balfanz, Herzog, & MacIver, 2007; Bowers, 2010b; Dynarski, 2004; Gleason & Dynarski, 2002)? Much of the research on students who fail to graduate high school has centered on a single dropout category

(Balfanz, et al., 2007), however recent research indicates that there may be multiple subgroups, or typologies of dropouts (Janosz, Archambault, Morizot, & Pagani, 2008). In addition, while many covariates associated with a student's increased risk of dropping out are well known, these covariates have historically only been estimated on the single dropout category. At the student level, these covariates include low grades, negative behavior, and low SES, among many others (Rumberger, 2004). In addition, school process and structure variables such as school size and student-teacher ratio have also been identified as increasing the likelihood that students may drop out (Lee & Burkam, 2003; Leithwood & Jantzi, 2009; Rumberger & Palardy, 2005).

The purpose of this study is to bring together these different domains within the dropout literature and test a combined model of high school dropout. This model estimates the associated influence of both student and school variables on the likelihood of student inclusion in different subgroups using student trajectory in teacher-assigned grades through the first three semesters of high school in a large nationally representative sample (ELS:2002). The likelihood of dropping out of high school is then assessed given a student's grade trajectory pattern. Using growth mixture modeling, we test a mediated model of the likelihood of students dropping out of high school. In this model, rather than test the direct effects of covariates on the likelihood of dropping out, we test the effects of a set of covariates on students' growth or decline achievement trajectories during the first three semesters of high school, which then influences the likelihood of students dropping out before the end of grade 12. In this way, this study helps to build and test a more complete model of student dropout that includes multiple typologies of students, based on student longitudinal achievement trajectories.

Dropout Typologies

While the majority of the research on dropouts considers students who leave school as a single "dropout" category (Balfanz, et al., 2007; Lee & Burkam, 2003; Rumberger & Palardy, 2005), an emerging research area focuses on describing what has been termed as "dropout typologies" (Bowers, 2010a; Fortin, Marcotte, Potvin, Royer, & Joly, 2006; Janosz, et al., 2008; Janosz, LeBlanc, Boulerice, & Tremblay, 2000; Lessard et al., 2008). This work acknowledges that rather than a single monolithic "dropout" category, student decisions to leave school are much more individualized and specific to certain subgroups within the dropout category. This dropout typology work has focused on identifying patterns of student engagement, behavior, and achievement to describe the different types of student processes that lead to dropping out.

To date, typology studies generally separate student dropouts into four distinct groups based on the characteristics students share with other dropouts. These groups are comprised of students who are either: *Disrupting School*, *Chronically Struggling with Academics*, *Bored with the Process*, or *Quiet* dropouts. Students

¹ This document is a preprint of an article published in 2012 in the Journal of Educational Research. Recommended Citation: Bowers, A.J., Sprott, R. (2012) Examining the Multiple Trajectories Associated with Dropping Out of High School: A Growth Mixture Model Analysis. *The Journal of Educational Research*, 105(3), 176-195. [doi:10.1080/00220671.2011.552075](https://doi.org/10.1080/00220671.2011.552075)

² Teachers College, Columbia University; Bowers@tc.edu; 525 W. 120th Street, New York, New York 10027.

ORCID: 0000-0002-5140-6428 ResearcherID: C-1557-2013

Note: This document was last updated on September 19, 2013

Disrupting School exhibit low grades and behavior problems. Students *Chronically Struggling with Academics* do not show behavior problems but demonstrate very low academic achievement. Those *Bored with the Process* have the highest grades of the dropouts but are disengaged from the process and show low commitment. Finally, *Quiet* dropouts appear most similar to graduates, but they do not possess the necessary support systems to persist in school when faced with certain obstacles. These students are particularly susceptible to rapid decline due to outside factors (Fortin, et al., 2006; Janosz, et al., 2000). Since typologies group potential dropouts by their scholastic needs, this approach provides school leaders with a critical tool. Using typologies, administrators, who have the unique ability to look across school data (Bowers, 2008, 2009) not only can identify potential dropouts, but typological categories can help school leaders guide the appropriate use of efficient dropout intervention strategies. However, while informative, the typology research to date has been limited, either to descriptive studies of a small sample of schools (Fortin, et al., 2006; Lessard, et al., 2008), or a focus exclusively on datasets of French Canadian students (Archambault, Janosz, Fallu, & Pagani, 2009; Janosz, et al., 2008) and has focused on using measurements of student engagement and behavior to identify the different dropout typologies.

In contrast to this typology literature, much of the dropout research over the past 40 years has considered student dropout as a single category. In many of these studies, a set of covariates are tested to estimate the effects on the likelihood of students dropping out of school (Balfanz, et al., 2007; Barrington & Hendricks, 1989; Lee & Burkam, 2003; Pallas, 1989; Rumberger, 1987, 2004; Rumberger & Palardy, 2005). These covariates include multiple categories such as background variables (such as student sex, SES, ethnicity, and family structure), student behavior and performance variables (extracurricular activities, discipline records, attendance, grades, test scores, grade retention), and school variables (student-teacher ratio, urbanicity, school size). Furthermore, rather than a single event, student dropout has come to be conceptualized as a “life course” perspective (Alexander, Entwisle, & Kabbani, 2001; Bowers, 2010a, 2010b; Entwisle, 1990; Finn, 1989; Pallas, 1989; Rumberger & Palardy, 2005), in which students experience a long history of challenges longitudinally with school that eventually leads to their decline and disengagement from the schooling process. However, while the longitudinal perspective is an improvement over a cross-sectional view of dropouts, for the majority of the empirical longitudinal studies on the life course, these studies have conceived of the dropout category as a single category and estimated the direct effects of the many covariates on dropping out. More recently, this type of research has been critiqued as missing as many as half of the students who are most likely to drop out and as mis-identifying students at risk of dropping out who actually graduate (Balfanz, et al., 2007; Gleason & Dynarski, 2002). We aim to address this issue by bringing together these different domains in the dropout literature by considering a mediated model of dropout. Rather than a single category, we test a model in which there may be different typologies of dropouts identified through longitudinal growth or decline in achievement, controlling for the many known covariates of dropping out.

Modeling Heterogeneous Growth Trajectories

The underlying issue at the heart of the past dropout typology work has been to find and describe homogenous subgroups of students who behave in similar ways that in some manner help predict if a student will graduate high school or not. This is difficult however, given that student data exists in large heterogeneous samples that vary across students, schools and over time. To address these issues of examining large heterogeneous longitudinal datasets that contain multiple homogenous subgroups, a form of multi-level structural equation modeling (SEM) known as Growth Mixture Modeling (GMM) has recently gained popularity (Dolan, 2009; Duncan, Duncan, & Strycker, 2006; Jung & Wickrama, 2008; Kaplan, Kim, & Kim, 2009; B. O. Muthén, 2004; B. O. Muthén et al., 2002; Nagin, 2005). GMM is a non-linear hierarchical modeling technique, similar to SEM, that allows for the identification of empirically defined subgroups in large longitudinal datasets while testing for the associated effects of a selection of variables at multiple levels within the model (Enders & Tofighi, 2008; Hix-Small, Duncan, Duncan, & Okut, 2004; Jung & Wickrama, 2008; B. O. Muthén, 2004; B. O. Muthén, et al., 2002; B. O. Muthén & Muthén, 2000; Vermunt, Tran, & Magidson, 2008; Wang & Bodner, 2007). These subgroups are generally referred to as latent trajectory classes. For dropout research, the use of GMM returns to the issue of dropout typologies. The vast majority of dropout studies to date that use multiple, hierarchical or logistic regression estimate a single homogenous growth trajectory for a dataset, taking as an assumption that the estimated parameters influence a single dataset-wide student trajectory (B. O. Muthén, 2004). However, as demonstrated in the dropout typology research, at least two sub-categories exist within any dropout dataset, students who graduate and those who do not. In addition, multiple sub-categories of students may also exist within the dropout group (Archambault, et al., 2009; Bowers, 2010a; Fortin, et al., 2006; Janosz, et al., 2008; Janosz, et al., 2000). Two recent research studies suggest that GMM may provide a useful means to not only help identify which students are most at risk of dropping out, but also estimate and control for the influence of specific variables on the likelihood that different subgroups may drop out.

In a study by Janosz et al. (2008), the researchers used GMM to study the relationship of student engagement to the risk of dropping out of high school for 13,300 French-Canadian high school students from low socioeconomic high schools across Quebec. Students were given an 18-item survey to assess student engagement. The authors used GMM to model the distribution of students by school engagement into two different latent class subgroups: 1) students with high initial and longer-term engagement from grades 7 through 12 and 2) students in a heterogeneous engagement trajectory group. Over 98% of the students in the first group graduated. The second group contained subgroups of heterogeneous latent classes with different initial engagement and long-term trajectories of engagement throughout high school, in which some subgroups increased engagement over the years, while others decreased. However, none in this second group were stable and high. These more heterogeneous engagement subgroups described 79% of the students who dropped out. However, many non-dropouts also had similar engagement patterns, such that only 27% of the students in the heterogeneous set of engagement trajectory subgroups dropped out. As hypothesized by Janosz et al. (2008), many of the students with low engagement with school through time may be members of the disengaged dropout typology.

The Janosz et al. (2008) study is an interesting example of the application of GMM to identifying more accurately the students most at risk of dropping out. However, the study is problematic for two main reasons. First, while the sample size was large, it is an intact sample of Quebec students from low SES schools, limiting the ability to generalize to other populations. Second, the use of assessments of student engagement, while interesting, is difficult to replicate, especially if the ultimate goal of at-risk identification research is application of the findings by teachers and school administrators. While theories of student engagement have been linked to understanding dropout risk for some time (Finn, 1989; Ream & Rumberger, 2008), additional surveys of student behavior are in addition to the many different types of data already collected in schools. To provide school leaders with accurate and useful information on who is most likely to drop out, researchers should strive to use the data already collected on students to aid in application and usefulness of the research findings (Bowers, 2009, 2010b; Catterall, 1998).

In a different study, as a means to demonstrate the usefulness of growth mixture modeling, Muthén (2004) presented a GMM analysis of the different longitudinal latent growth curve trajectories (latent classes) of 3,102 students from the Longitudinal Study of Youth (LSAY) from 1987. Muthén states that the objective of GMM is to empirically identify unobserved latent classes within a dataset, also referred to above as typologies, clusters, and subgroups (these terms are used interchangeably here). GMM provides a means to estimate the probability that an individual is a member of any one specific unobserved subgroup while estimating and controlling for the associated influence of covariates (Duncan, et al., 2006; Jung & Wickrama, 2008; B. O. Muthén, 2004). The study modeled the different longitudinal mathematics standardized test score trajectories of students in the LSAY sample over grades 7, 8, 9 and 10, estimating the outcome of the likelihood that the students dropped out of high school by grade 12, controlling for multiple covariates such as gender, ethnicity, mother's education, as well as overall school poverty and student-teacher ratio (B. O. Muthén, 2004). The study found three latent classes, two that corresponded to graduation and one to dropout. In the two graduation subgroups, mathematics achievement began relatively high and continued high throughout high school. In the dropout subgroup, identified in the study as similar to the disengaged dropout typology from the literature discussed above, students began with lower mathematics achievement and their achievement increased slower than the other two groups, or decreased. Sixty-nine percent of the students in the low subgroup dropped out, demonstrating one of the most accurate dropout identification and description methods to date.

While Muthén (2004) did provide a powerful and useful means of identifying and studying different subgroups of dropouts from large datasets using GMM, there are three main issues with the study. First, the analysis was presented as a tutorial for GMM, not as a study of dropout typologies and predictors. Little of the dropout literature is referenced, and the final estimated model is not focused on testing and extending the past research. Rather, Muthén presented it as an interesting example of the usefulness of GMM. Second, because the study was a tutorial of GMM, the model is underparameterized and does not test specific research questions. Third, the study focused on the use of LSAY data, that while interesting, is somewhat dated since it does not represent a more current dataset with additional variables. Despite these issues, GMM provides a new and powerful means to explore and

test the theories currently proposed in the above reviewed dropout literature, controlling for the longitudinal nature of a dropout dataset, and allowing for the empirical identification of subgroups of dropouts, while appropriately controlling for multiple known covariates.

A Combined Model of Typologies and Dropout

The issue we consider here is a combined model that reconsiders the single dropout category as possibly including multiple subgroups of dropouts that align with the dropout typology literature. In addition, we aim to estimate the influence of the main covariates associated historically with an increased likelihood of students dropping out of high school on each of the different identified subgroups. Therefore, for the first time in the dropout literature this study brings together and integrates three main topics: dropout typologies, student growth in achievement, and dropout covariates. To this end, this study has three main research questions.

- 1) To what extent do students who drop out of high school pattern into different subgroups based on their trajectory in teacher assigned grades?
- 2) To what extent do the dropout typologies identified in the literature relate to different subgroups of dropouts identified by achievement trajectories?
- 3) To what extent are past dropout covariates associated with different dropout achievement trajectories?

METHODS

Sample

This study is a secondary analysis of the Education Longitudinal Study 2002 (ELS:2002) (NCES, n.d.). ELS:2002 is a nationally-representative dataset for a sample of students collected across multiple years by the U.S. National Center for Education Statistics (NCES). Approximately 15,400 grade 10 students were sampled in 2002 and then again in 2004 and 2006 (Bozick et al., 2006; Ingles et al., 2004; Ingles et al., 2007). A multitude of variables were collected at all time-points, including all student teacher-assigned grades in all courses in grades 9 through 12, if students had graduated or had dropped out, and a self-survey about their behavior and aspirations. The ELS:2002 provides a unique opportunity to analyze a deep and nationally-representative dataset focused on the high school years. However, the dataset is similar to the well-known High School & Beyond study (Ingles, et al., 2004; Rasinski, Ingels, Rock, Pollack, & Wu, 1993) in that it is somewhat limited in scope since it focuses solely on the final years of voluntary enrollment in secondary school. Due to the requirements of the analytic model discussed below, we restricted the analysis to an ELS:2002 subsample, namely students only from public schools (private schools excluded) and only students from schools on a semester or quarter system. Thus, the final sample size for the study was $n = 5400$ students. Due to confidentiality requirements for the dataset, some numbers have been rounded.

Teacher Assigned Grades

An extensive set of literature has identified teacher assigned grades as one of the most well-known predictors of which students are most likely to drop out of school (Allensworth & Easton, 2005; Balfanz, et al., 2007; Barrington & Hendricks,

1989; Battin-Pearson et al., 2000; Bowers, 2010b; Kirschenbaum, Napier, & Simon, 1971; Lloyd, 1978; Rumberger, 1987, 2004; Wehlage, Rutter, Smith, Lesko, & Fernandez, 1989). For this study, we modeled longitudinal growth or decline in teacher assigned grades from grades 9 through 10 to describe different student achievement trajectories. The ELS:2002 dataset provides a unique opportunity to examine the longitudinal change in student grades since it also contains a high school transcript study (Bozick, et al., 2006). Students were recruited and included in ELS:2002 by NCES when the students were in grade 10 in 2002. In 2004, during the first follow-up (F1), the entire high school transcripts were collected of all students in the study. This included each teacher assigned grade for every course for all students. Grades were recorded as letter grades and converted to a standard five-point scale (0-4) (Bozick, et al., 2006). To replicate past research that has identified non-cumulative Grade Point Average (GPA) as superior to a cumulative GPA calculation (Bowers, 2009, 2010b), we calculated non-cumulative GPA for the first and second semesters of grade 9 and the first semester of grade 10 by calculating the arithmetic mean grade point from all recorded grades at each of these three time points for each student. To also include students from schools on a quarter term system in the final analytic model, the first two quarters of a school year were considered as a single semester in the GPA calculations, and the second two quarters were considered the second semester in an academic year.

Variables Included in the Analytic Model

We aim to replicate, extend and integrate the past research on the associated influence of multiple variables on different subgroups of students who graduate and drop out. We based our decisions to include variables in the analytic model on literature that has previously shown significant associated effects in either dropout identification, prediction or the dropout typology domains. In addition to teacher assigned grades discussed above, three types of variables were included in the model, including 1) student and school background and demographics, 2) student behaviors, and 3) school structure variables. Descriptive statistics for each of the variables as well as item loadings for the combined scale are provided in Appendix 1 and 2.

Student and School Background and Demographics:

A long history of research has indicated that males and students from historically disadvantaged backgrounds are more likely to drop out of school. This includes *African American* and *Hispanic* students as well as students from low socio-economic (*SES*) backgrounds and *non-traditional families* (less than two parents or guardians in the home) (Cataldi, et al., 2009; Laird, Cataldi, Ramani, & Chapman, 2008; B. O. Muthén, 2004; Ream & Rumberger, 2008; Rumberger, 2004; Rumberger & Palardy, 2005). School background variables historically associated with dropping out include school locale (*urban, rural*, with suburban as the reference group) as well as the average poverty level of the student body, usually measured by the *percent of students in a school receiving free or reduced priced lunches* (B. O. Muthén, 2004; Rumberger, 2004; Rumberger & Palardy, 2005). See Appendix 1 for the complete list of variables used in the study linked to ELS:2002 variables codes.

Student Behavior Variables

A range of in-school student behaviors have historically been associated with students who drop out of school. One of the strongest factors associated with student dropout is grade

retention (Bowers, 2010b; Jimerson, Anderson, & Whipple, 2002; Roderick, 1994). In addition, student engagement with school, as measured through the number of student *extracurricular* activities, is a well-known factor associated with dropping out (Broh, 2002; Finn, 1989; Mahoney, 2000; Rumberger, 1987, 2004) (see Appendix 1). Student misbehavior and disengagement with school can be measured through the number of student suspensions and punishments and student tardiness, truancy and overall delinquency (Balfanz, et al., 2007; Fortin, et al., 2006; Gaviria & Raphael, 2001; Janosz, et al., 2008; Rumberger, 2004) designated here as a composite measure of *negative behavior* (see Appendix 2).

School Variables

The emerging literature on school variables has nominated multiple variables that may be associated with students who drop out. School enrollment is well known as being associated with dropping out, in which *large schools* and *extra large schools* (more than 1200 students) experience higher rates of student drop out (Lee & Burkam, 2003; Leithwood & Jantzi, 2009; Rumberger, 2004; Rumberger & Palardy, 2005) while *small schools* and medium sized schools experience lower rates. Here medium sized schools were the reference group (see Appendix 1). School size categorization was based on the categories used in previous studies (Lee & Burkam, 2003; Leithwood & Jantzi, 2009; Rumberger, 2004; Rumberger & Palardy, 2005). School processes associated with dropping out include the *student-teacher ratio* (here grand mean centered), and the level of *academic press* of the school (Lee & Burkam, 2003; Lillard & DeCicca, 2001; McNeil, Coppola, & Radigan, 2008; Rumberger & Palardy, 2005).

Dropout

The final type of data included in the analysis was graduation or dropout status, measured as the ELS:2002 variable F2EVERDO. This dropout indicator defines student dropout in the following ways (Bozick, Lauff, & Wirt, 2007; Ingles, et al., 2007). The present study is concerned with graduation from high school with a full high school diploma or the failure to graduate, termed in much of the research as “dropping out”. However, the opportunities presented to students in the United States who do not wish to complete high school in four years are many. This complicates these two categories of graduation or dropout. One of the most problematic issues with longitudinal studies of student dropout has been transfer out of the school district under study, thus making the status of student high school graduation unknown (Bowers, 2007, 2009, 2010b). Fortunately with the ELS:2002 dataset, students were followed if they transferred out of a district to their new district, making loss of data due to transfer status less of a problem than with past studies. A second issue with dropout status are students who do not graduate on time in the traditional four years, yet obtain a full high school diploma within four and a half, five or more years (Balfanz, et al., 2007). The longitudinal nature of the ELS:2002 sample collection, with the second follow-up occurring two years after student on-time graduation from high school in 2006, helps address this issue, since students who did not graduate on time yet did eventually graduate up to two years late are included in the dataset. A third and final issue with dropout status are students who opt to take an exam and receive a G.E.D. (general educational development) rather than complete the traditional requirements for a regular high school diploma. Past research on the G.E.D. option has indicated that a G.E.D. is not equivalent to a regular high school diploma (Cameron & Heckman, 1993;

Tyler, 2003). Because the proposed study here is focused on graduation from high school with a regular high school diploma in a traditional high school curriculum, and the G.E.D. is not considered equivalent, students who received a G.E.D. rather than a regular high school diploma were considered as having dropped out. Thus, the ELS:2002 variable F2EVERDO was used as the dropout indicator assessing if a student ever dropped out by the second follow-up at F2 in 2006.

Analytic Model

To assess the extent to which student longitudinal trajectories in non-cumulative GPA are associated with dropping out of high school, we used Growth Mixture Modeling (GMM) as the primary means of data analysis. GMM has recently emerged as a powerful and effective means to empirically identify subgroups within datasets (Duncan, et al., 2006; Jung & Wickrama, 2008; B. O. Muthén, 2004; Nagin, 2005; Wang & Bodner, 2007). As discussed above, in education GMM has been used to describe dropout engagement typologies from French Canadian student datasets (Janosz, et al., 2008), as well identify a dropout category from among multiple high school student mathematics achievement growth trajectories (B. O. Muthén, 2004). GMM is a non-linear multilevel modeling technique, similar to structural equation modeling (SEM) and hierarchical linear growth modeling (HLM) that allows for the identification of empirically defined homogenous subgroups in large heterogeneous datasets while testing for the associated effects of a selection of variables at multiple levels within the model. We used a single nested model analyzing growth in non-cumulative GPA to identify patterns associated with different GPA trajectories of students more or less likely to drop out. Here, students are nested in time, and then student longitudinal GPA trajectories are nested within four latent trajectory classes. These latent class GPA trajectories are akin to different typologies of students, but rather than defined by a set of engagement variables, as in Janosz (2008), the typology, or “latent trajectory class” in GMM terms, is defined by student longitudinal growth or decline in non-cumulative GPA. This method: a) appropriately controls for the longitudinal nature of the schooling and dropout process through examining change in GPA over time; b) identifies the extent to which graduates and dropouts in the sample pattern into different GPA trajectories; and c) assesses the extent to which the different trajectories are associated with covariates as well as the different intercepts and slopes within each latent trajectory class.

GMM is a person-centered statistic that appropriately controls for the conditional nature of a dataset such as ELS:2002. For a review of growth mixture modeling, please see Muthén (2004). Briefly, past methods of analyzing multilevel achievement growth trajectories, such as HLM, assume that the growth trajectory is either fixed, or homogenous for a dataset (B. O. Muthén, 2004; Raudenbush & Bryk, 2002). In contrast, GMM takes a structural equation modeling (SEM) approach in which the growth mixture is a multinomial logistic regression in which individual growth trajectories through time empirically define different latent trajectory classes of individuals, controlling for a set of covariates, and estimating the probability of experiencing a distal outcome given a specific latent class trajectory.

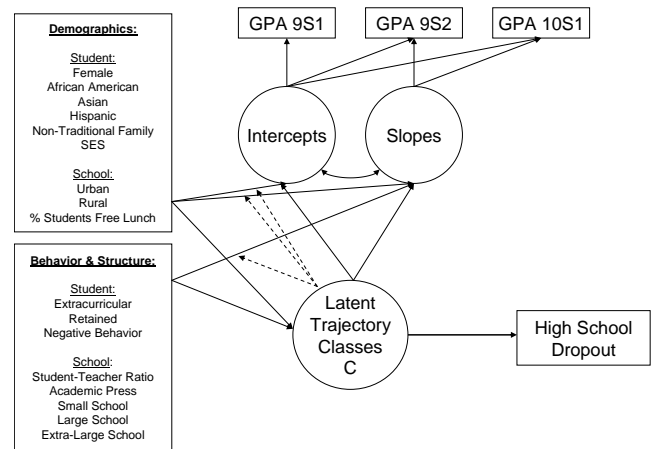


Figure 1: Growth mixture model for the simultaneous estimation of latent trajectory classes using non-cumulative GPA from the first three semesters of high school.

Following the recommendations from the GMM literature (Duncan, et al., 2006; Jung & Wickrama, 2008; B. O. Muthén, 2004; Wang & Bodner, 2007), Figure 1 represents the GMM for this study by adapting the nomenclature from SEM for representing the overall model. Similar to HLM, GMM estimates the slope of the intercepts as well as the slope of the growth trajectory slopes from a set of longitudinal achievement scores, here using non-cumulative GPA from grade 9 semester 1 and 2 and grade 10 semester 1 (Figure 1, upper set of boxes and the intercept and slopes circles). Traditional HLM stops here, estimating a single homogenous slope of the intercepts and slope of the overall growth trajectory. In other words, the entire dataset is fit to a single growth trajectory line, and then the associated influence of covariates is assessed. However, the question of interest here is if growth trajectories in GPA (the “growth” part of a growth mixture model) are heterogeneous with multiple homogenous subgroups (the “mixture” part of a growth mixture model). If multiple subgroups in longitudinal growth in GPA exist, different covariates may be differentially associated with the different latent growth trajectories. In short, GMM is designed such that the technique can empirically assess if the traditional dropout category is a single category or contains multiple subgroups, basing the subgroups on different overall student growth or decline in GPA through the first three semesters of high school. GMM is designed to address these issues through identifying the subgroups as “latent trajectory classes” (B. O. Muthén, 2004). As depicted with the arrows extending from the latent trajectory class circle to the intercepts and slopes in Figure 1, GMM can identify the latent classes within a dataset, each of which may have a different set of intercepts and growth trajectories through the longitudinal achievement data. Covariates may also vary differently with each of the latent classes as well as on the intercepts and slopes for each latent class, represented in Figure 1 as arrows extending from the covariate boxes on the left towards the circles. In addition, GMM can estimate the likelihood of a distal outcome, such as dropping out of school, for each of the latent classes associated with the different significant achievement growth trajectories in the dataset, represented here as the high school dropout box on the right. The effects of each of the covariates on each latent trajectory class were allowed to vary in the model on the intercepts and slopes, as represented by the dotted arrows in Figure 1. Additionally, the intercepts and slopes for each latent trajectory class were correlated within the model,

represented by the curved arrow in Figure 1. The analysis was conducted using MPLUS 5.21 (L. K. Muthén & Muthén, 2007).

Thus, Figure 1 depicts the growth mixture model using ELS:2002 data to identify significant student GPA trajectories associated with different probabilities of graduating or dropping out of school. The GMM is structured in the following five ways. First, achievement growth trajectories through the first three semesters of high school are estimated using non-cumulative GPA. As discussed above, teacher assigned grades appear to be one of the most important variables in identifying which students are most at risk of dropping out. Here, non-cumulative GPA was used to estimate growth curves across the first three semesters of high school for all students in the dataset.

Second, the ELS:2002 transcript study data provides a unique opportunity to study trajectories of student achievement associated with dropout or graduation due to the structure of the survey and data collection. Dropout research using achievement data is inherently difficult, since if a student drops out, then their data history ends much earlier than graduates do. This creates a missing data problem. To estimate growth curves for any fitted model for SEM and GMM, the data must be complete (Duncan, et al., 2006). For the students in ELS:2002 in grades 11 and 12, if a student dropped out, and thus has missing GPA data, that data is not missing at random since it is inherently linked to the issue that the student dropped out. Fortunately, the inherent design of ELS:2002 data collection can be taken advantage of to address this point. The base year for ELS:2002 was 2002 when the students were in grade 10. Students were then tracked over the following years. While all dropout events were captured from grade 10 semester 2 through grade 12, students who dropped out prior to grade 10 are not part of the dataset. This inherently focuses any study using ELS:2002 on the later years of high school, which is of interest here given the above discussion of student dropout. Although the beginning of time for the study was grade 10 semester 2, transcripts for all students were collected from grades 9 through 12. Therefore, grading data for all students present in the base year extended back in time to three semesters before the start of the study. We took advantage of this design issue of ELS:2002 that teacher assigned grades were present in the dataset for students three semesters before the beginning of the survey. As depicted in Figure 1, growth curves were estimated using the three semesters for which grades were recorded before students began to drop out of school in the dataset. This eliminates the missing data problem with achievement data for students dropping out after data collection began. Additionally, this design is supported by the past GMM study that demonstrated that significant growth trajectories can be estimated using three time points previous to the end of grade 10 (B. O. Muthén, 2004).

The third structural component of the model is the covariates. Controlling for specific known covariates associated with the outcome helps to estimate the appropriate latent class growth curves in the GMM (Duncan, et al., 2006; B. O. Muthén, 2004). Here, the multiple known covariates with dropout discussed above were included in the model. Two sets of covariates were included, those that conceptually could affect the intercepts and the slopes (such as background variables) and those that assessed variables during grades 9 and 10, and so could only affect the slopes but not the intercepts. The different sets of arrows from the covariates box on the left of Figure 1 indicate these differences in the model.

The fourth structural component of the model concerns the study's central question of reconsidering the historical single dropout category as instead containing multiple typologies of dropouts. In terms of the model depicted in Figure 1, past OLS and HLM regression studies have studied the direct effects of the covariates on the left with high school dropout on the right, without the growth model in the middle. Here, we reconsider the direct effects model of dropout, and instead estimate a mediated model with a latent trajectory class mediating variable placed between the covariates and dropout. Using this mediated model, we propose and test here, that rather than directly affecting the probability of dropping out, these covariates may instead influence a student's probability of inclusion in a growth or decline achievement trajectory, which then influences a student's likelihood of graduating or dropping out of high school. In this way, the mediated model brings together the multiple domains of the current research on dropping out, including the multiple covariates, longitudinal growth or decline in achievement, and the multiple typologies of dropping out. Unfortunately, while we would have liked to also include the direct effects of the covariates on dropout in the model, doing so would preclude our ability to estimate the probability of student inclusion in each latent class, since such a model is not identified.

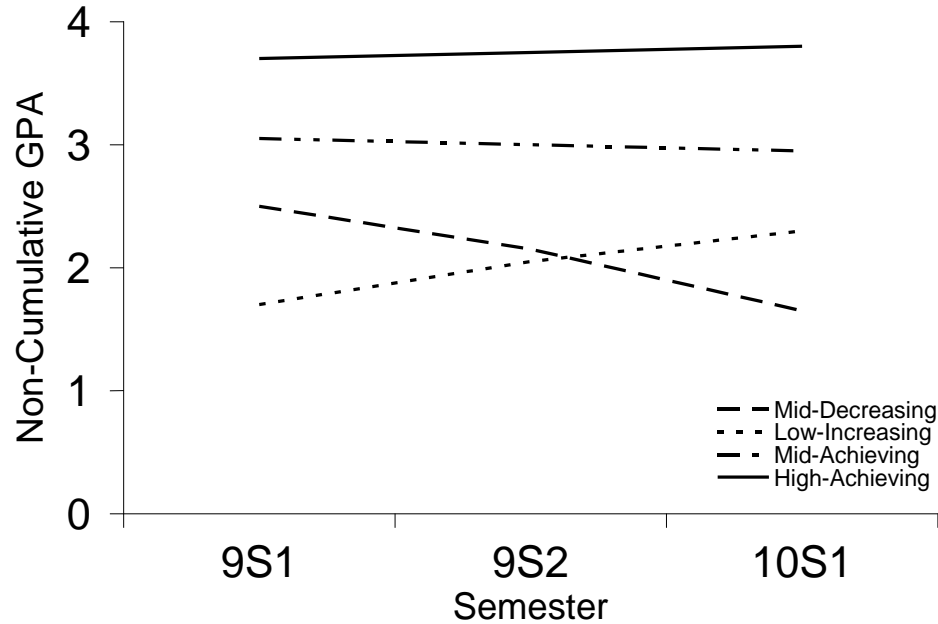
The final structural component of the model is the use of the main survey component of ELS:2002 in the base year. In grade 10 semester 2, students were surveyed on a variety of issues pertaining to school climate, discipline, how much they liked school, and their expectations (Ingles, et al., 2007). For the present study, students in the different latent trajectory classes identified in the GMM were matched to their survey responses in an attempt to describe the latent class trajectories. Two independent sample t-tests assuming unequal variances were used to examine differences in the means of responses between the different latent classes. Cohen's *d* was calculated using the pooled variances.

RESULTS

This study examines the intersection of three domains within the high school dropout literature. First, we reconsider the homogenous dropout category as a heterogeneous category potentially including multiple homogenous subgroups. Second, because the literature indicates that teacher assigned grades are highly predictive of dropping out, we aim to identify these dropout subgroups using student non-cumulative GPA trajectories during the first three semesters of high school. Third, we estimate the associated influence of both student and school covariates on the likelihood of membership in different subgroups. Our hypothesis is that rather than a single category, different types of student dropouts exist within schools, that these different subgroups can be identified using non-cumulative GPA achievement trajectories, and that both student and school characteristics influence the likelihood that students pattern into trajectories either headed for graduation or increased chances of dropping out.

We estimated a four latent class trajectory solution using the GMM diagramed in Figure 1. Previous research using high school test score data from the U.S. in the 1980s identified three latent classes, two of which were middle or high achieving groups, with one low achieving group containing the majority of the dropouts (B. O. Muthén, 2004). To replicate and extend this research to

A:



B:

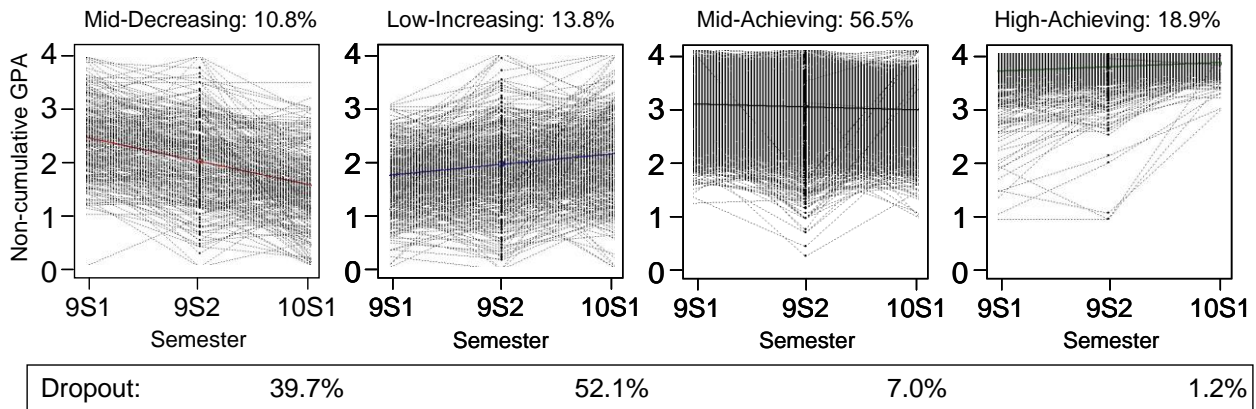


Figure 2: Longitudinal Non-Cumulative GPA Trajectories in the First Three Semesters of High School. **A)** Four latent class growth or decline trajectories were extracted from the growth mixture model based on student patterns of non-cumulative GPA during grade 9 semester 1 and 2 and grade 10 semester 1. Sample means for each latent class are plotted. **B)** Latent class trajectory model estimated means are plotted (bold lines) with each individual student's GPA pattern for each latent class. Students were surveyed in grade 10 semester two and their progress tracked over the subsequent years, including if they dropped out of high school. The four class growth mixture model identified two classes that included 24.6% of the sample with an increased likelihood of dropping out of high school, accounting for 91.8% of the dropouts.

identify multiple potential dropout categories, we increased the number of potential latent class trajectories to four. A major critique of GMM is that the number of latent classes to be estimated is selected *a priori*, based on the literature and theory, much like SEM (B. O. Muthén, 2004; Nylund, Asparouhov, & Muthén, 2007). Caution must be used in estimating too many trajectories, since power to estimate each class and the effects of the covariates decreases substantially with each additional latent class estimated (Bauer & Curran, 2003). Thus, we took a conservative view and estimated a four-class model with separate

intercepts and slopes for each latent trajectory class in non-cumulative GPA over time.

The four class GMM is a good solution that fits the data well. Model fit was evaluated following the recommended procedures for GMM (Jung & Wickrama, 2008; Nylund, et al., 2007) using the log likelihood H_0 value, BIC value, and entropy estimate. The final four-class model resulted in a log likelihood H_0 value of -9212.024, a BIC value of 19928.154, and an entropy estimate of 0.631 which compares favorably to past GMM dropout studies

Table 1: Multinomial logistic regression model estimation of the likelihood of latent class trajectory categorization in comparison to Mid-Achieving as the reference group.

	Mid-Decreasing		Low-Increasing		High-Achieving	
	Coefficient	Odds	Coefficient	Odds	Coefficient	Odds
Student background						
Female	-0.239		-0.644~	0.525	0.601**	1.824
African American	0.494		-0.132		-2.906**	0.055
Asian	-0.329		-0.033		0.458~	1.581
Hispanic	1.001*	2.721	0.065		-0.337	
Non-traditional family	0.340		0.392		0.219	
SES	-0.959*	0.383	-1.220*	0.295	0.641***	1.848
Student behaviors						
Extracurricular	-0.117		-0.121**	0.886	0.045***	1.046
Retained	1.699***	5.468	1.746***	5.732	-0.843	
Negative behavior	2.741***	15.502	2.632***	13.902	-1.966***	0.140
School structure						
Urban	0.194		-0.495		-0.357	
Rural	0.084		0.681~	1.976	-0.014	
% Free lunch	0.011		0.006		0.001	
Student-teacher ratio	-0.002		0.033~	1.034	0.014	
Academic press	0.080		0.063		0.119	
Small school	0.565		-0.338		0.053	
Large school	0.732		-0.094		0.059	
Extra-large school	0.889~	2.433	-0.378		-0.544*	0.580
Odds of dropping out for category versus Mid-Achieving		43.478		44.260		0.550

Note: Parameter estimates and odds ratios for each respective latent class are in comparison to the normative reference class Mid-Achieving as a function of the covariates.

Note: SES = socio-economic status

Note: ~p<0.1, *p<0.05, **p<0.01, ***p<0.001

with entropy values between 0.6 and 0.7 (Janosz, et al., 2008; B. O. Muthén, 2004). The fit of the four-class solution versus the $k=1$ three-class solution was evaluated using the recommended fit statistics of the Lo-Mendell-Rubin likelihood ratio test (LMR) and the parametric Bootstrapped Likelihood Ratio Test (BLRT) using the recommended 100 bootstrap draws (Jung & Wickrama, 2008; L. K. Muthén & Muthén, 2007; Nylund, et al., 2007). The LMR p-value equaled 0.348 while the BLRT p-value equaled p<0.001. As recommended by Nylund et al. (2007), we considered the BLRT as the more accurate and robust test of the correct number of classes and thus consider the four-class growth mixture model as fitting the data well.

The three main types of results from the GMM are presented in Figure 2, Table 1 and Table 3. The GMM simultaneously estimates a multinomial logistic regression, estimating inclusion in four separate latent trajectory classes based on four separate growth (or decline) regression estimates, each with its own intercept and slope factors, controlling for and estimating the influence of the covariates identified in the model in Figure 1. The odds of dropping out of high school were estimated using these four trajectories.

The four latent trajectory classes in non-cumulative GPA identified by the model from grade 9 semester 1 through grade 10 semester 1 are presented in Figure 2. Sample means for the four trajectories are plotted together in Figure 2A. The four trajectories are designated as Mid-Decreasing, Low-Increasing, Mid-Achieving and High-Achieving based on the grade 9 semester 1 intercepts and the direction of the slopes through the three semesters. Figure 2B plots each of the 5,400 student's actual

trajectories for the entire dataset in non-cumulative GPA for the first three semesters of high school, with the model estimated means for each trajectory in bold. As presented in Figure 2, the Mid-Decreasing and Low-Increasing trajectories account for 24.6% of the sample (10.8% and 13.8% respectively) and 91.8% of the dropouts (39.7% and 52.1% respectively). This is in comparison to the Mid-Achieving and High-Achieving trajectories that together accounted for 75.4% of the sample but only 8.2% of the dropouts. Thus, the GMM identified four latent trajectory classes based on student growth or decline in teacher assigned grades within the first three semesters of high school. For students who failed to complete high school, we identified two different types of dropouts as belonging to either the Mid-Decreasing or Low-Increasing trajectory groups. In addressing the first research question, our data suggests that rather than a single homogenous category, dropout is heterogeneous with students who drop out belonging to at least two homogenous subgroups of GPA trajectory. This is in comparison to past studies that have considered dropout as a single homogenous category.

Table 1 presents the multinomial logistic regression model estimates for the covariates for the four latent class trajectories. Here, the likelihood of inclusion in each group is compared to inclusion in the Mid-Achieving trajectory group. We selected the Mid-Achieving trajectory group as the normative reference group because the Mid-Achieving trajectory represents the majority of the students (56.5% of the sample) and appears to also represent the normative latent trajectory class of a mid-achieving student who graduates high school on time. Thus, the three other trajectories are compared to Mid-Achieving as the reference

Table 2: Variable means and standard deviations, disaggregated by latent trajectory class

	Mid-Decreasing		Low-Increasing		Mid-Achieving		High-Achieving	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD
GPA9S1	2.45	0.698	1.61	0.565	2.97	0.570	3.66	0.425
GPA9S2	2.05	0.781	1.86	0.741	2.88	0.627	3.73	0.343
GPA10S1	1.41	0.687	2.09	0.796	2.85	0.539	3.82	0.180
Female	0.47	0.499	0.36	0.481	0.50	0.500	0.67	0.472
African American	0.35	0.477	0.14	0.352	0.13	0.345	0.01	0.068
Asian	0.07	0.258	0.07	0.255	0.10	0.295	0.18	0.387
Hispanic	0.22	0.417	0.07	0.252	0.06	0.231	0.03	0.164
Non-traditional family	0.37	0.483	0.36	0.480	0.19	0.391	0.14	0.348
SES	-0.16	0.513	-0.28	0.467	0.16	0.590	0.47	0.627
Extracurricular	2.16	4.091	1.90	3.939	5.33	5.895	7.81	6.191
Retained	0.28	0.450	0.35	0.477	0.06	0.241	0.01	0.122
Negative behavior	1.12	0.684	1.02	0.613	0.48	0.317	0.30	0.244
Urban	0.42	0.495	0.11	0.318	0.21	0.405	0.14	0.348
Rural	0.14	0.342	0.40	0.491	0.27	0.446	0.28	0.447
% Free lunch	27.93	19.388	22.12	16.773	18.24	14.973	15.64	12.448
Student-teacher ratio	0.82	5.560	0.13	5.572	-0.22	3.659	0.30	3.632
Academic press	2.98	0.813	2.90	0.796	3.01	0.785	3.12	0.751
Small school	0.15	0.362	0.26	0.439	0.21	0.410	0.26	0.437
Large school	0.32	0.466	0.22	0.412	0.24	0.427	0.28	0.450
Extra-large school	0.38	0.487	0.13	0.332	0.24	0.429	0.19	0.388

group in Table 1. Each coefficient represents the logit estimate of the influence of the covariate on the likelihood of inclusion in each of the three latent trajectories in comparison to having been included in the Mid-Achieving trajectory. Significant coefficients were converted to odds (e^{logit}) to aid in interpretation. As an example from Table 1, students who were retained at any time previous to grade 10 were 5.468 times more likely to be in the Mid-Decreasing trajectory than the Mid-Achieving trajectory, confirming and replicating the extensive past research on the negative impact of retention on student achievement and dropout reviewed above. Variables are grouped according to student background, student behaviors and school structure.

Overall, student inclusion in the *Mid-Decreasing* trajectory appears to be influenced the most if they were (+)Hispanic and from low (-)SES background. In addition, these students were (+)retained much more often than Mid-Achieving students and had significantly higher levels of (+)negative behavior. Mid-Decreasing students also attended schools more often with (+)extra large enrollments. Significant coefficients for the *Low-Increasing* category were (-)female, (-)SES, (-)extracurricular, (+)retained, (+)negative behavior, (+)rural, and (+)student-teacher ratio. Significant coefficients for the *High-Achieving* category were (+)female, (-)African American, (+)Asian, (+)SES, (+)extracurricular, (-)negative behavior, and (-)extra-large schools. In addition, students in the Mid-Decreasing and Low-Increasing groups were 43 and 44 times more likely to dropout, respectively, than students in the Mid-Achieving group (Table 1, bottom row).

Thus, for the two trajectories associated with dropping out of high school, these findings replicate and extend much of the literature on the variables most associated with dropping out of school, but with two important advances. First, rather than estimating the direct effect of each covariate on dropping out, the model estimated here is more akin to a structural equation model, estimating the effect of each covariate on the mediating latent class trajectory variable (Figure 1), which our hypothesis posits lies between the covariates and dropping out. Second, we have

identified two trajectories that are associated with a higher likelihood of dropping out, with somewhat different patterns of significant covariates. Both the Mid-Decreasing and Low-Increasing trajectories include students who were retained and engaged in higher amounts of negative behavior. However, the likelihood of inclusion in Low-Increasing is dependent upon being male (a negative female coefficient) while Mid-Decreasing does not appear to depend on gender. A similar difference was observed with the Low-Increasing students having participated in fewer hours of extracurricular activity per week than Mid-Decreasing students. Additionally, Mid-Decreasing inclusion is significantly related to school size, while Low-Increasing is not.

Table 2 presents the means and standard deviations for each variable for each of the four identified latent trajectory classes. While Table 2 provides a means to examine the differences and similarities across each identified group, for brevity, we focus here on the differences between the Mid-Decreasing and Low-Increasing groups. The Mid-Decreasing and Low-Increasing students had the highest chances of dropping out. Examining the differences between these two groups of students reveals that Mid-Decreasing students had mean GPAs starting in grade 9 semester 1 of 2.45, and decreasing to 2.05 in grade 9 semester 2, and 1.41 in grade 10 semester 1. They were more often African American (35% versus 14%) or Hispanic (22% versus 7%) and were retained somewhat less than Low-Increasing students (28% versus 35%). In addition, Mid-Decreasing students were much more often from large or extra large urban schools with higher percentages of free lunch students. This is in comparison to Low-Increasing students who had mean GPAs in grade 9 semester 1 of 1.61, rising to 1.86 in grade 9 semester 2 and 2.09 in grade 10 semester 2. Low-Increasing students were more likely to be male and from rural schools. Interestingly, there were few differences between these two groups in a variety of other well-known dropout predictors, such as being from a non-traditional family, low SES, negative behavior, and school academic press. Next, we turn to analyzing the intercepts and slopes of the growth portion of the growth mixture model.

Table 3: Multivariate regression estimates on the intercepts and slopes for each of the identified non-cumulative GPA GMM latent trajectory classes

	Mid-Decreasing	Low-Increasing	Mid-Achieving	High-Achieving
Intercept	2.446	1.705	3.009	3.643
Female	0.017	-0.027	0.103**	0.055~
African American	-0.189	0.035	-0.102*	-0.324***
Asian	0.184*	0.040	0.069*	-0.013
Hispanic	-0.073	-0.089	-0.034	0.012
Non-traditional family	-0.079	-0.030	-0.061	-0.013
SES	0.154*	-0.076	0.094*	0.088**
Urban	-0.195~	0.001	0.063	0.029
Rural	0.078	0.156	0.013	-0.008
R-Square	0.205	0.046	0.048	0.118
Slope	-0.464	0.227	-0.056	0.080
Student background				
Female	-0.013	0.496*	0.011	-0.056
African American	-0.124	-0.111	0.110	0.406***
Asian	-0.415***	0.009	-0.100~	0.045
Hispanic	0.248	0.244	0.032	0.004
Non-traditional family	0.008	-0.042	-0.003	-0.037
SES	-0.109	0.010	-0.016	-0.100*
Student behaviors				
Extracurricular	-0.004	0.105	0.034	-0.048~
Retained	0.130	-0.135	-0.006	-0.025
Negative Behavior	0.240~	-0.151	-0.155*	0.045
School structure				
Urban	0.274	0.118	-0.020	-0.017
Rural	-0.229	-0.155	-0.021	-0.029
% free lunch	-0.080	0.185~	-0.017	0.058
Student-teacher ratio	0.122	-0.312***	-0.040	-0.098**
Academic press	0.090	0.151	0.008	0.002
Small school	0.526~	-0.177	0.014	-0.125**
Large school	0.232	-0.123	-0.031	-0.127***
Extra-large school	0.345	0.179	0.056	-0.063~
R-square	0.623	0.592	0.056	0.216

Note: Coefficients are expressed as effect sizes as a change in Y standard deviation units for a 1 standard deviation change in X.

Note: SES = socio-economic status

Note: ~p<0.1, *p<0.05, **p<0.01, ***p<0.001

Table 3 presents the multiple regression estimate effect sizes for the intercepts, slopes and covariates for each of the four extracted latent trajectory classes. The data in Table 3 provide a portrait of the most significant variables influencing GPA in the first semester of high school (the intercepts), as well as student GPA trajectory through the first three semesters of high school (the slopes). As noted in the GMM in Figure 1, covariates were split into two types, those that were student, school and community background variables and were regressed on both the intercepts and slopes, and behavior and school structure variables that were regressed only on the slopes. One of the main advantages of growth mixture modeling is that the model simultaneously estimates the intercepts and slopes for each separate latent trajectory class, here four different regression models. Since the latent trajectory class extraction is based on the growth or decline curve part of the model, each set of intercepts and slopes should

be considered significantly different from each of the other trajectory groups.

As opposed to standard OLS or HLM regressions, the intercepts for a GMM are of substantive interest since each of the latent trajectory classes has its own intercept and slope. Thus, different model estimated mean values for the intercepts can be interpreted as the mean GPA in grade 9 semester 1 for the average student in each of the four trajectories groups. In addition, the GMM estimates the influence of the covariates on the intercepts, and estimates the amount of the variance explained by the included covariates. As shown in Table 3 (top section), while 20.5% of the variance was explained in the Mid-Decreasing intercept (R-square=0.205) and 11.8% of the variance was explained in the High-Achieving intercepts, the covariates did not explain more than 5% of the variance in intercepts of the other two trajectory groups. Hence, while about 20% of the variance in grade 9

Table 4: Mean differences between Mid-Decreasing and Low-Increasing trajectories in survey item responses.

Questions from ELS:2002 Base Year Survey	Grand Mean	Mid-Decreasing Mean (SD)	Low-Increasing Mean (SD)	t Value (df)	p-value (Cohen's d)
Teachers are interested in students	2.83	2.63 (0.81)	2.69 (0.76)	-1.292 (1105)	0.197 (-0.08)
Teachers praise effort	2.73	2.58 (0.77)	2.64 (0.78)	-1.288 (1121)	0.198 (-0.08)
In class often feels put down by teachers	1.85	2.06 (0.78)	1.97 (0.73)	1.956 (1120)	0.051 ~ (0.12)
Does not feel safe at this school	1.73	1.93 (0.84)	1.81 (0.76)	2.311 (1106)	0.021 ** (0.15)
Disruptions often get in the way of learning	2.46	2.51 (0.86)	2.47 (0.89)	0.86 (1119)	0.390 (0.05)
There are gangs in this school	2.11	2.42 (0.99)	2.14 (0.90)	4.847 (1104)	<0.001 *** (0.30)
Classes are interesting and challenging	2.57	2.41 (0.78)	2.32 (0.77)	1.898 (1127)	0.058 ~ (0.12)
Satisfied by doing what is expected in class	2.63	2.54 (0.79)	2.40 (0.77)	2.853 (1122)	0.004 ** (0.18)
Has nothing better to do than school	2.21	2.19 (0.84)	2.34 (0.86)	-2.898 (1121)	0.004 ** (-0.18)
Education is important to get a job later	3.58	3.42 (0.72)	3.38 (0.71)	0.917 (1123)	0.359 (0.06)
School is a place to meet friends	3.06	2.80 (0.79)	2.97 (0.73)	-3.768 (1124)	<0.001 *** (-0.22)
Plays on a team or belongs to a club	2.62	2.20 (0.88)	2.17 (0.90)	0.512 (1117)	0.610 (0.03)
Teachers expect success in school	2.67	2.55 (0.86)	2.44 (0.84)	2.075 (1125)	0.038 * (0.13)
Parents expect success in school	3.46	3.41 (0.74)	3.30 (0.74)	2.423 (1128)	0.016 * (0.15)
How much likes school	2.11	1.96 (0.57)	1.88 (0.61)	2.121 (1088)	0.034 * (0.13)
How far in school student thinks will get	5.26	4.58 (1.67)	4.13 (1.64)	4.105 (923)	<0.001 *** (0.27)
Plans to continue education after high school	4.56	4.34 (1.01)	4.15 (1.11)	2.54 (870)	0.011 * (0.18)

Note: All variables were student self reported and all but the last two were coded on a 1-4 point Likert scale where 1=*strongly disagree*, 2=*disagree*, 3=*agree*, 4=*strongly agree*.

Note: Item "how far in school student thinks will get" was coded as 1=*less than high school*, 2=*high school graduation or GED*, 3= *attend or complete 2-year college*, 4=*attend college*, 5=*graduate from college*, 6=*obtain masters*, 7=*obtain a doctoral degree*.

Note: Item "plans to continue education after high school" was coded as 1=*no, don't plan to continue education*, 2=*yes, but don't know when*, 3=*yes, after out of high school over 1 year*, 4=*yes, after out of high school 1 year*, 5=*yes, right after high school*.

Note: ~p<0.1, *p<0.05, **p<0.01, ***p<0.001

semester 1 GPA was explained in the Mid-Decreasing group by the covariates (+)Asian, (+)SES and (-)urban, none of the covariates were significant on the intercepts in the Low-Increasing trajectory. This may indicate that how high or low students start in high school GPA in the Low-Increasing trajectory has more to do with variables at the earlier levels of schooling that were not available to run in the model than with the background variables included. In other words, we were unable to explain the variance in the intercepts for the Low-Increasing students using the included background predictor variables. Therefore, since their grade 9 semester 1 GPAs are on average the lowest of all three groups, one interpretation is that these students may have entered high school on a low-grade trajectory, and thus variables that would help to explain this beginning GPA would be found at earlier grade levels, such as elementary and middle school. To a lesser extent, this same difference in intercepts was also apparent in comparing the High-Achieving and Mid-Achieving trajectories.

The bottom half of Table 3 presents the slope coefficients for the four latent class trajectory models. About 60% of the variance in the non-cumulative GPA slope across the three semesters was explained by the model for the Mid-Decreasing and Low-Increasing trajectories (Table 3, bottom row). Only 5.6% of the variance was explained in the Mid-Achieving slope and 21.6% of the variance in the High-Achieving slope. Of substantive interest here are the differences in sign and significance across the covariates for the Mid-Decreasing and Low-Increasing trajectories. However, interpretation of the sign of the slope coefficients is not straightforward, since the slope for Mid-Decreasing is negative (-0.464).

As with any regression coefficient, the sign of a significant coefficient relates to a change in the magnitude of the slope. A positive coefficient will increase the magnitude of a slope no matter if the slope is positive or negative, while a negative coefficient will decrease the magnitude of a slope. Thus, a *positive* coefficient on a *negative* slope indicates that the slope becomes *more negative* as the variable increases while a *negative* coefficient on a *negative* slope decreases the magnitude of the slope, making the slope *less negative*. In examining the significant coefficients on the slope of the Mid-Decreasing group (Table 3, lower portion), (-)Asian indicates that for Asian students in the group, their downward trajectory in GPA across the three semesters is significantly less steep on average than the rest of the group. Students with more negative behavior (+) and from smaller schools (+) experience more steep average declines in GPA in the Mid-Decreasing group. The slope coefficients for Low-Increasing (with a positive overall slope) differed from those of the other latent groups with (+)females and students in schools with higher (+) percentages of free lunch students rising faster, while students from schools with higher (-)student-teacher ratios rose slower through time in their GPA as would be expected given previous literature (Table 3, lower portion). Additionally, the High-Achieving coefficients are of interest. Increasing (-)student-teacher ratios had a negative influence on the slope, as did (-)small, (-)large and (-)extra large schools, indicating that students in the High-Achieving trajectory that were from medium sized schools (the school size reference group) with the lowest student-teacher ratios had the highest slopes. Also, while African American was not significantly related to the intercepts and slopes for Mid-Decreasing and Low-Increasing, the model estimates suggest that African American students in the High-Achieving trajectory started in grade 9 semester 1 significantly

below the average (negative coefficient on the intercepts), but then rose significantly faster than the average (positive coefficient on the slope). However, as noted in Table 2, few African-American students patterned into the High-Achieving trajectory.

We turn now to examining the differences between the Mid-Decreasing and Low-Increasing latent trajectory classes. Together, Figure 2, and Tables 1, 2 and 3 present an interesting picture of two very different types of students, both of which drop out at substantial rates. Both groups of students participated in extracurricular activities much less than the rest of the sample, and they were retained and demonstrated negative behaviors on average more than the remaining student population. The Mid-Decreasing students appeared to start with a mid-GPA on average in grade 9, but then fell over time. These students were from large schools with higher numbers of Hispanic students and low SES students. In contrast, the Low-Increasing students appeared to start with a low GPA, and rise moderately through time. These students were much more likely to be male, but their likelihood of inclusion in the latent trajectory and their slope were not significantly influenced by school size, indicating that the other covariates may be explaining this variance, such as SES, gender and rural.

Comparing the Mid-Decreasing and Low-Increasing Groups

To further explore and describe the differences in the Mid-Decreasing and Low-Increasing latent trajectory classes, student responses to additional items from the ELS:2002 survey were analyzed using independent comparison t-tests (see Table 4). Table 4 presents the mean response, standard deviations, significance and effect size to each question for the two different groups. The sample grand mean was included to provide a means to determine if the responses of either group were above or below the average response, but was not included in the significance test. Students were given the survey in grade 10 semester 2. Items included in Table 4 relate to the school climate, including discipline, safety, academic press, student behavior and student expectations. The Mid-Decreasing group appeared to have significant mean differences around issues of safety, discipline, friendships and academic press. This may reflect many of the differences found in the GMM, with Mid-Decreasing students attending larger more urban schools and here reporting much more often that the student “does not feel safe in this school” and “there are gangs in this school” and agreeing less with the question “school is a place to meet friends”. Mid-Decreasing students also indicated that they felt that they would get farther in school in comparison to the Low-Increasing students and attend post-secondary school sooner in response to the questions “how far in school student thinks will get” and “plans to continue education after high school”. However, both groups felt that they would complete less schooling than the average student in the sample.

Conversely, the Low-Increasing group had significant mean differences on the items related to interest and satisfaction in school and school expectations. Low-Increasing students disagreed on average more with the question “classes are interesting and challenging”, “satisfied by doing what is expected in class”, “teachers expect success in school” and “parents expect success in school” in comparison with the Mid-Decreasing group. Low-Increasing students also agreed more that the student “has nothing better to do than school” but, interestingly, reported that

they liked school less in response to the question “how much likes school”.

DISCUSSION

Research on dropping out of high school using large representative datasets has historically considered dropout as a single category. However, the more descriptive and qualitative dropout typology literature suggests that dropout may be a heterogeneous category with multiple homogeneous subgroups. In this study, we show evidence that supports the multiple dropout typology literature, arguing for a more complex view of the dropout process, in that we have identified two different student achievement trajectories associated with dropping out of high school. Rather than the theory tested in previous literature of the direct effects of covariates on the likelihood of dropping out, we proposed and tested a mediated variable theory of covariates affecting C (inclusion in and trajectory of four different latent class trajectories) which then leads to dropping out, where C is a latent variable that mediates the effects of covariates on a student’s likelihood of dropping out. As with SEM, the aim is to test theory by proposing a model and estimating the extent to which it fits or does not fit the data (Duncan, et al., 2006; Jung & Wickrama, 2008; Kaplan, et al., 2009; B. O. Muthén, 2004; Wang & Bodner, 2007). For dropout research, this effort helps to build actionable theory about a student’s likelihood of dropping out in an effort to inform future interventions and school practice in attempting to reduce a school’s dropout rate and help students graduate. We acknowledge that this study provides only preliminary evidence in support of an expanded theory of the dropout process. Nevertheless, using this tentative expanded model, we argue that researchers and practitioners interested in interrupting the dropout process may want to move their focus back one step from students dropping out of school, and instead focus on shifting a student’s achievement trajectory from one associated with dropping out, such as the Mid-Decreasing or Low-Increasing groups, to the Mid-Achieving or High-Achieving trajectories associated with graduation.

Expected versus Unexpected Dropouts & Dropout Typologies

Rather than a single dropout category, two distinct groups associated with dropping out of high school emerged in our data. Here we identified the Mid-Decreasing and Low-Increasing trajectories in which the Mid-Decreasing students had higher initial grades followed by a steep decline while the Low-Increasing students’ grades started lower and increased slightly. While both groups displayed many of the previously identified predictors of dropping out, we posit that the Mid-Decreasing and Low-Increasing student latent class trajectories represent the difference between what can be considered as “expected” versus “unexpected” dropout typologies.

Low-Increasing

The Low-Increasing group’s average grades started at the lowest level of any of the four groups in our data set. The position of their initial GPA, along with their grade 10 responses regarding low scholastic expectations possibly speaks to a history of academic difficulties and frustration. Previous literature contains several typologies that show similarities to the low-increasing group. Identified previously as a School and Social Adjustment Difficulties type (Fortin et al., 2006), Low-Achiever Pushouts (Kronick & Hargis, 1998), or the Maladjusted and Low-Achievers (Janosz, 2000), this typology of students eventually tire

of the difficult educational processes encountered throughout their time in the system and either feel pushed out of school due to misbehavior and low grades or they opt for what they perceive as easier alternatives such as the pursuit of a GED. When compared to the Mid-Decreasing group, the Low-Increasing group exhibited significantly higher levels of frustration with school. They reported that parent and teacher expectations were lower, that they liked school the least, and compared to the Mid-Decreasing students, the Low-Increasing students expressed lower expectations of continuing with school.

These Low-Increasing students are the types of students schools may expect to dropout. They conform to many of the past dropout prediction variables, such as their negative behavior, low grades, and the point that they are more often male. All of these characteristics put them in an at-risk category previously identified as highly associated with dropping out. Indeed, this finding mirrors much of the recent dropout identification literature (Allensworth & Easton, 2007; Balfanz, et al., 2007; Bowers, 2010b; Rumberger & Palardy, 2005). However, as noted by Gleason and Dynarski (2002), many of these types of dropout predictors mis-identify students. Many times, almost half of the students identified as potential dropouts never dropout, whereas half of the students who do dropout were never identified as at-risk of dropping out. This problem was acknowledged by Balfanz et. al. (2007), in that while they were able to accurately identify 60% of the dropouts using these types of expected dropout variables (low grades, high negative behavior), they were unable to accurately identify early about 40% of the students who eventually dropped out. Here, we identified 52.1% of the students who dropped out as Low-Increasing students, students that we propose as the expected dropouts are similar to the classic “at-risk” dropout category. In comparison, we nominate the Mid-Decreasing group, which contained 40% of the dropouts in the sample, as an “unexpected” dropout category.

Mid-Decreasing

In contrast to the typologies of students who struggle with school, the Mid-Decreasing group reflects students who may be a more unexpected type of dropout. The Mid-Decreasing students started with higher average grades but their non-cumulative GPA declined over time. This decline late in the school process (grades 9 and 10) suggests that their difficulty with school was not as chronic as it was for the Low-Increasing group. Instead, in many ways the Mid-Decreasing students appear to be somewhat unaware that they are on a trajectory associated with increased chances of dropping out. In comparison to the Low-Increasing, Mid-Decreasing are almost evenly divided between females and males, and student school expectations are much closer to the average student. However, these students are significantly less likely to engage in extracurricular activities. Moreover, our data showed that the Mid-Decreasing students’ attend schools in areas of social turmoil in large or extra-large urban low-SES schools.

Similar to Kronick and Hargis’ (1998) Quiet dropouts, Janosz et al. (2000) Quiets, and the Anti-Social Coverts described by Fortin et al., (2006), the Mid-Decreasing students likely persist in school only until outside demands become too great to resist. Thus, this group’s dropping out can be difficult to predict using in-school variables since their decline may happen rapidly and they may not have a history of low grades until the semesters right before they drop out. As one hypothesis, these students attend larger urban schools and may be lost in the shuffle, disengaged from school not because they are experiencing difficulties with academics, but

because they are not engaged in the social aspects of schooling. This idea is supported by these students' lower response to the question "school is a place to meet friends". In comparison to the Low-Increasing students that are more likely struggling with the academics of school, the Mid-Decreasing students' decline in grades may not be an indication of their lack of ability with academics, but rather an indication of their increasing disengagement with schooling. Thus, the Mid-Decreasing student is an unexpected dropout, a type of student that schools previously have had little information on to help them address the specific needs of these students.

Implications of a Multiple Category View of Dropouts

The main finding, that dropout involves multiple and distinct achievement trajectory groups that can be identified using non-cumulative GPA, is important since an improved understanding of dropout typologies could enable schools to provide individual students with more effective dropout interventions. To date, most dropout intervention studies have shown little effect on helping to prevent student dropout (Dynarski et al., 2008; Dynarski, 2004; Dynarski & Gleason, 2002). However, this point returns to the issue of mis-identification of students most likely to dropout. As noted above, the literature to date on the "flags" most associated with students dropping out of school has either not identified or mis-identified between 40% and 50% of the students who eventually dropout (Balfanz, et al., 2007; Bowers, 2010b; Gleason & Dynarski, 2002). Following the recommendations of the student engagement literature (Archambault, et al., 2009; Finn, 1989; Fredricks, Blumenfeld, & Paris, 2004; Janosz, et al., 2008), many recent dropout prevention efforts attempt to increase student engagement with school, matching students with school mentors, involving students with in-school and after school programs, and attempting to engage a student in the social aspects of schooling, all with limited results (Dynarski, et al., 2008; Dynarski, 2004). However, the students identified as at-risk, are identified using the classic dropout identification flags, here shown to be associated with the Low-Increasing expected dropout group, students that we propose are chronically struggling with school. Our results suggest that the lack of an effect of engagement interventions may be because the interventions are provided to the expected dropout, not to the unexpected Mid-Decreasing dropout who we hypothesize may benefit the most from these types of school re-engagement interventions. As noted above, Mid-Decreasing students appear to disengage with high school fairly rapidly, as evidenced through their average decline in grades. Using a multi-category view of dropouts, our results suggest that Low-Increasing students may benefit the most through academic interventions, while Mid-Decreasing students may benefit the most through engagement interventions. To date, the Mid-Decreasing students have historically gone un-identified as at-risk of dropping out. The results presented here indicate that schools may be able to identify Mid-Decreasing students as students with declining non-cumulative GPAs.

Growth Model Trajectories

While the mixture part of the GMM discussed above identified the different latent trajectory classes, the growth part of the model also provides interesting information on these different types of students by examining both the intercepts and longitudinal slopes in non-cumulative GPA. For the intercepts, our model explained about 20% of the variance in grade 9 semester 1 GPA for the Mid-Decreasing students, 11% of the variance for the High-Achieving students, but less than 5% of the variance in the intercepts for the Low-Increasing and Mid-Achieving students

(see Table 3). For the intercepts part of the model, only background variables were included. For the Mid-Decreasing students, and to a lesser extent, the High-Achieving students, non-school variables (the background variables) had a significant influence on their beginning high school grades. Stated another way, for the Mid-Decreasing students, at least one fifth (20%) of the variance in the average grade 9 semester 1 GPA was not due to the school. Conversely, for the Low-Increasing students, none of the background variables were significant. One interpretation of these results is that because background variables did not explain a large portion of the variance in the beginning grades for two of the trajectories, the intercepts for these two trajectories (Low-Increasing and Mid-Achieving) may be the most influenced by in-school variables that were not included in the model or that occurred earlier in a student's career in elementary or middle school. This interpretation is supported by the dropout literature on the life course perspective (Alexander, et al., 2001; Entwisle, 1990; Finn, 1989; Pallas, 1989), in which for the majority of the students who drop out, they experience continual challenges throughout their schooling process, which eventually leads to their disengagement and leaving school. In many ways, the Low-Increasing "expected" dropout students described here fit into this life course perspective. Indeed, the life course perspective posits that the school is the main influence on these students (rather than family or outside influences), and this is supported here by the significant negative impact of the student-teacher ratio variable in the slopes part of the model for the Low-Increasing students. Additionally, while we can only make a tentative claim at this point, it may be that the Low-Increasing students are also the students most affected by the well studied problematic transition from grade 8 to grade 9 (Allensworth & Easton, 2007; Belcher & Hatley, 1994; Benner & Graham, 2009; Bowers, 2010b; Rumberger, 1995; Zvoch, 2006), since they begin with the lowest average grades of the different trajectories. In contrast, in-line with the hypothesis that the Mid-Decreasing students are the "unexpected" dropouts that past dropout identification methods have failed to identify early as at risk of dropping out, a significant portion of the variance in the intercepts (20%) is explained by the background variables. This indicates that for these students, the school may have less of an effect on their beginning grades in high school, and this finding of few significant school variables is repeated in the declining achievement slope for the Mid-Decreasing students.

Non-cumulative GPA

Our results support and extend the rich set of literature that indicates that teacher assigned grades are highly predictive of students dropping out. In this study, we used the longitudinal change in non-cumulative GPA in the growth-model part of the growth mixture model to help define the different latent trajectories mixtures. Recent research has indicated that rather than use cumulative GPA, which is problematic as a variable due to the cumulative dependent nature of how it is calculated, student non-cumulative GPA reveals the semester-to-semester ups and downs in grades which can be used to identify students most at risk of dropping out (Bowers, 2010b). Additionally, while much of the literature on teacher-assigned grades has argued that grades are a weak assessment of academic knowledge, emerging research has demonstrated that teacher assigned grades are a useful multi-dimensional assessment of both student academic knowledge and a student's ability to negotiate the social processes of school (Bowers, 2009). This dual-assessment nature of grades comes from teachers assigning grades based on both academic achievement as well as a variety of student behaviors,

such as participation, attendance and behavior. It has been argued that grades are predictive of student dropout because they incorporate these behavior aspects, which are an evaluation of if a student is able to negotiate the social processes of school, which if they are successful, leads to graduation (Bowers, 2009).

In this study we have replicated and extended this work, demonstrating that not only do low grades identify the majority of students who will drop out, but that trajectories of GPA, rather than single grades or course failures, can inform an expanded theory of the dropout process. Here, the Low-Increasing or expected dropouts started high school with low grades. These students may be identified using many of the “dropout flags” nominated in the literature, from low grades and course failures (Allensworth & Easton, 2005, 2007; Balfanz, et al., 2007; Bowers, 2010b; Hickman, Heinrich, Bartholomew, & Mathwig, 2008). Nevertheless, these students account for only about half of the dropouts in the sample. The Mid-Decreasing, or unexpected dropouts include 40% of the dropouts. These students started with fairly high grades, but then declined over time. Our hypothesis here is that it is these unexpected dropouts that have gone unidentified in the past longitudinal dropout identification research. We posit that since this subgroup of students starts high school with mean grades of a B- or C+, and that they participate in school extracurricular activities much less than the average student, that the decline in grades over the first three semesters of high school goes unnoticed by the schools. Combined with the emerging literature that grades may be a multi-dimensional assessment of both academic knowledge and student behaviors, it may be that the decline in grades by Mid-Decreasing students is not an indication of academic problems, but rather an indication that Mid-Decreasing students are disengaging with the schooling process.

Limitations

Results and interpretations of this study are limited in three main ways. First, because of the design and timing of the ELS:2002 survey, this study is limited to the final three years of high school. Students who dropped out prior to grade 10 semester 2 were not included in the dataset. This is problematic because student dropout prior to grade 10 is well known, especially at the grade 8 to grade 9 transition (Benner & Graham, 2009; Bowers, 2010b; Cohen & Smerdon, 2009; Rumberger, 1995; Zvoch, 2006). Thus, the findings of this study cannot be interpreted as referring to trends that may exist for all dropouts, but rather as a description of what may be occurring for dropouts in the final three years of high school. Second, a major limitation is sample size. While this study uses one of the largest sample sizes in the dropout literature, outside of the studies done with large urban cohorts such as with Chicago or Philadelphia (Allensworth & Easton, 2005; Balfanz, et al., 2007), it is a subsample of the full ELS:2002 dataset. As noted above, the growth mixture model requires complete data on all of the covariates, which limited the sample size to 5,400. While this reduction in sample size was unavoidable, it means that we are unable to generalize to the entire ELS:2002 dataset, and thus must consider this study more as an initial descriptive study of the different latent trajectories in need of further replication, rather than an inferential study.

The third major limitation is an inherent issue with the *a priori* specification of the number of latent trajectory classes in the growth mixture model. As noted above, the number of latent classes for a GMM is stated prior to running a model and is based on theory and the literature. However, this *a priori* specification

of the number of latent classes is problematic because the GMM bases the estimation of the fit of the number of latent classes on the nonnormality of the data (Bauer & Curran, 2003). As an example, if a sample contains multiple modes, and theory indicates that these modes are true homogenous latent classes in a heterogeneous population, then GMM is designed to detect these different subgroups based on this non-normality. However, as stated by Bauer and Curran (2003), the question for a researcher is “do the components represent true latent subgroups in the population, or are they serving only to approximate what is in fact a homogenous but non-normal distribution?” (p.343). They also provide evidence in the same study that over-extraction of latent classes (estimating more classes than actually exist in the population) dramatically reduces the power to detect significant parameters throughout the model (Bauer & Curran, 2003). For these reasons, as noted above in the analytic model, rather than conduct an exhaustive search of model fit with increasing numbers of latent classes, we erred on the conservative side and selected to estimate a GMM with one additional latent class (four classes) than the only previous achievement and dropout GMM study to date (B. O. Muthén, 2004). Thus, our model may underestimate the true number of latent trajectory classes, but reduces the chances of a type I error while advancing this research domain. Future studies will work to further refine the model.

Conclusion

In conclusion, we see this study as a valuable contribution to the dropout literature, combining the to-date separate domains of the effects of grades and the covariates on different typologies of students most likely to drop out of school, rather than a single category of dropout. The findings of this study suggest that future dropout interventions should 1) track student longitudinal non-cumulative GPA as an early indicator of dropout risk; 2) devise interventions that address the different needs of the different typologies of high school dropouts; and 3) work to shift low achievement trajectories to the higher, more engaged achievement subgroups.

RECOMMENDED CITATION

Bowers, A.J., Spratt, R. (2012) Examining the Multiple Trajectories Associated with Dropping Out of High School: A Growth Mixture Model Analysis. *The Journal of Educational Research*, 105(3), 176-195. [doi:10.1080/00220671.2011.552075](https://doi.org/10.1080/00220671.2011.552075)

REFERENCES

- Alexander, K. L., Entwisle, D. R., & Kabbani, N. S. (2001). The dropout process in life course perspective: Early risk factors at home and school. *The Teachers College Record*, 103(5), 760-822.
- Allensworth, E. M., & Easton, J. Q. (2005). The on-track indicator as a predictor of High School graduation (Vol. 2006): Consortium on Chicago School Research at the University of Chicago.
- Allensworth, E. M., & Easton, J. Q. (2007). What matters for staying on-track and graduating in Chicago public high schools: A close look at course grades, failures, and attendance in the freshman year. Chicago: Consortium on Chicago School Research.
- Archambault, I., Janosz, M., Fallu, J.-S., & Pagani, L. S. (2009). Student engagement and its relationship with early high school dropout. *Journal of Adolescence*, 32, 651-670.

- Balfanz, R., Herzog, L., & MacIver, D. J. (2007). Preventing student disengagement and keeping students on the graduation path in urban middle-grades schools: Identification and effective interventions. *Educational Psychologist, 42*(4), 223-235.
- Barrington, B. L., & Hendricks, B. (1989). Differentiating characteristics of high school graduates, dropouts, and nongraduates. *Journal of Educational Research, 82*(6), 309-319.
- Battin-Pearson, S., Abbott, R. D., Hill, K. G., Catalano, R. F., Hawkins, J. D., & Newcomb, M. D. (2000). Predictors of early high school dropout: A test of five theories. *Journal of Educational Psychology, 92*(3), 568-582.
- Bauer, D. J., & Curran, P. J. (2003). Distributional assumptions of growth mixture models: Implications for overextraction of latent trajectory classes. *Psychological Methods, 8*(3), 338-363.
- Belcher, D. C., & Hatley, R. V. (1994). A dropout prediction model that highlights middle level variables. *Research in Middle Level Education, 18*(1), 67-78.
- Benner, A. D., & Graham, S. (2009). The transition to high school as a developmental process among multiethnic urban youth. *Child Development, 80*(2), 356-376.
- Bowers, A. J. (2007). *Grades and data driven decision making: Issues of variance and student patterns*. Ph.D., Michigan State University, East Lansing. <http://eric.ed.gov/?id=ED538574>
- Bowers, A. J. (2008). Promoting Excellence: Good to great, NYC's district 2, and the case of a high performing school district. *Leadership and Policy in Schools, 7*(2), 154-177. [doi:10.1080/15700760701681108](https://doi.org/10.1080/15700760701681108)
- Bowers, A. J. (2009). Reconsidering grades as data for decision making: More than just academic knowledge. *Journal of Educational Administration, 47*(5), 609-629. [doi:10.1108/09578230910981080](https://doi.org/10.1108/09578230910981080)
- Bowers, A. J. (2010a). Analyzing the longitudinal K-12 grading histories of entire cohorts of students: Grades, data driven decision making, dropping out and hierarchical cluster analysis. *Practical Assessment Research and Evaluation, 15*(7), 1-18. <http://pareonline.net/pdf/v15n7.pdf>
- Bowers, A. J. (2010b). Grades and Graduation: A Longitudinal Risk Perspective to Identify Student Dropouts. *Journal of Educational Research, 103*(3), 191-207. [doi:10.1080/00220670903382970](https://doi.org/10.1080/00220670903382970)
- Bozick, R., Lauff, E., & Wirt, J. (2007). *Education longitudinal study of 2002 (ELS:2002): A first look at the initial postsecondary experiences of the high school sophomore class of 2002*. Washington, D.C.: National Center for Education Statistics, Institute of Education Sciences, U.S. Department of Education.
- Bozick, R., Lyttle, T., Siegel, P. H., Ingels, S. J., Rogers, J. E., Lauff, E., & Planty, M. (2006). *Education Longitudinal Study of 2002: First follow-up transcript component data file documentation*. Washington, DC: National Center for Education Statistics, U.S. Department of Education.
- Broh, B. A. (2002). Linking extracurricular programming to academic achievement: Who benefits and why? *Sociology of Education, 75*(1), 69-91.
- Cameron, S. V., & Heckman, J. J. (1993). The nonequivalence of high school equivalents. *Journal of Labor Economics, 11*(1), 1-47.
- Cataldi, E. F., Laird, J., & KewalRamani, A. (2009). *High school dropout and completion rates in the United States: 2007*. Washington, DC: National Center for Education Statistics, Institute of Education Sciences, U.S. Department of Education.
- Catterall, J. S. (1998). Risk and resilience in student transitions in high school. *American Journal of Education, 106*(2), 302-333.
- Cohen, J. S., & Smerdon, B. A. (2009). Tightening the dropout tourniquet: Easing the transition from middle to high school. *Preventing School Failure, 53*(3), 177-184.
- Dolan, C. V. (2009). Structural equation mixture modeling. In R. E. Millsap & A. Maydeu-Olivares (Eds.), *The SAGE handbook of quantitative methods in psychology* (pp. 568-591). Thousand Oaks, CA: SAGE Publications.
- Duncan, T. E., Duncan, S. C., & Strycker, L. A. (2006). *An introduction to latent variable growth curve modeling: Concepts, issues and applications* (2nd ed.). Mahwah, NJ: Lawrence Erlbaum Associates.
- Dynarski, M., Clarke, L., Cobb, B., Finn, J., Rumberger, R. W., Smink, J., . . . Gill, B. (2008). *Dropout prevention: A practice guide*. Washington, DC: National Center for Education Evaluation and Regional Assistance, Institute of Education Sciences, U.S. Department of Education.
- Dynarski, M. (2004). Interpreting the evidence from recent federal evaluations of dropout-prevention programs: The state of scientific research. In G. Orfield (Ed.), *Dropouts in America: Confronting the graduation rate crisis* (pp. 255-268). Cambridge, MA: Harvard Education Press.
- Dynarski, M., & Gleason, P. (2002). How can we help? What we have learned from recent federal dropout prevention evaluations. *Journal of Education for Students Placed at Risk, 2002*(1), 43-69.
- Enders, C. K., & Tofghi, D. (2008). The impact of misspecifying class-specific residual variances in growth mixture models. *Structural Equation Modeling: A Multidisciplinary Journal, 15*(1), 75-95. [doi: 10.1080/10705510701758281](https://doi.org/10.1080/10705510701758281)
- Entwisle, D. R. (1990). Schools and the adolescent. In S. S. Feldman & G. R. Elliott (Eds.), *At the threshold: The developing adolescent* (pp. 197-224). Cambridge, MA: Harvard University Press.
- Finn, J. D. (1989). Withdrawing from school. *Review of Educational Research, 59*(2), 117-142.
- Fortin, L., Marcotte, D., Potvin, P., Royer, E., & Joly, J. (2006). Typology of students at risk of dropping out of school: Description by personal, family and school factors. *European Journal of Psychology of Education, 21*(4), 363-383.
- Fredricks, J. A., Blumenfeld, P. C., & Paris, A. H. (2004). School engagement: Potential of the concept, state of the evidence. *Review of Educational Research, 74*(1), 59-109.
- Gaviria, A., & Raphael, S. (2001). School-based peer effects and juvenile behavior. *The Review of Economics and Statistics, 83*(2), 257-268.
- Gleason, P., & Dynarski, M. (2002). Do we know whom to serve? Issues in using risk factors to identify dropouts. *Journal of Education for Students Placed at Risk, 7*(1), 25-41.
- Hickman, G. P., Heinrich, R. S., Bartholomew, M., & Mathwig, J. (2008). Differential developmental pathways of high school dropouts and graduates. *Journal of Educational Research, 102*(1), 3-14.
- Hix-Small, H., Duncan, T. E., Duncan, S. C., & Okut, H. (2004). A multivariate associative finite growth mixture modeling approach examining adolescent alcohol and marijuana use. *Journal of Psychopathology and Behavioral Assessment, 26*(4), 255-270. [doi: 10.1023/B:JOBA.0000045341.56296.f](https://doi.org/10.1023/B:JOBA.0000045341.56296.f)

- Ingles, S. J., Pratt, D. J., Rogers, J. E., Siegel, P. H., Stutts, E. S., & Owings, J. A. (2004). *Education longitudinal study of 2002: Base year data file user's manual*. Washington, D.C.: National Center for Education Statistics, Institute of Education Sciences, U.S. Department of Education.
- Ingles, S. J., Pratt, D. J., Wilson, D., Burns, L. J., Currivan, D., Rogers, J. E., & Hubbard-Bednasz, S. (2007). *Education longitudinal study of 2002: Base-year to second follow-up data file documentation*. Washington, DC: National Center for Education Statistics, Institute of Education Sciences, U.S. Department of Education.
- Janosz, M., Archambault, I., Morizot, J., & Pagani, L. S. (2008). School engagement trajectories and their differential predictive relations. *Journal of Social Issues, 64*(1), 21-40.
- Janosz, M., LeBlanc, M., Boulerice, B., & Tremblay, R. E. (2000). Predicting different types of school dropouts: A typological approach with two longitudinal samples. *Journal of Educational Psychology, 92*(1), 171-190.
- Jimerson, S. R., Anderson, G. E., & Whipple, A. D. (2002). Winning the battle and losing the war: Examining the relation between grade retention and dropping out of high school. *Psychology in the schools, 39*(4), 441-457.
- Jung, T., & Wickrama, K. A. S. (2008). An introduction to latent class growth analysis and growth mixture modeling. *Social and Personality Psychology Compass, 2*(1), 302-317.
- Kaplan, D., Kim, J.-S., & Kim, S.-Y. (2009). Multilevel latent variable modeling: Current research and recent developments. In R. E. Millsap & A. Maydeu-Olivares (Eds.), *The SAGE handbook of quantitative methods in psychology* (pp. 592-612). Thousand Oaks, CA: SAGE Publications.
- Kirschenbaum, H., Napier, R., & Simon, S. B. (1971). *Wad-jaget? The grading game in American education*. New York City: Hart Publishing Company.
- Laird, J., Cataldi, E. F., Ramani, A. K., & Chapman, C. (2008). Dropout and completion rates in the United States: 2006. Washington, DC: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics.
- Lee, V. E., & Burkam, D. T. (2003). Dropping out of high school: The role of school organization and structure. *American Educational Research Journal, 40*(2), 353-393.
- Leithwood, K., & Jantzi, D. (2009). A review of empirical evidence about school size effects: A policy perspective. *Review of Educational Research, 79*(1), 464-490. doi: 10.3102/0034654308326158
- Lessard, A., Butler-Kisber, L., Fortin, L., Marcotte, D., Potvin, P., & Royer, E. (2008). Shades of disengagement: High school dropouts speak out. *Social Psychology of Education, 11*(1), 25-42.
- Lillard, D. R., & DeCicca, P. (2001). Higher standards, more dropouts? Evidence within and across time. *Economics of Education Review, 20*(5), 459-473.
- Lloyd, D. N. (1978). Prediction of school failure from third-grade data. *Educational and Psychological Measurement, 38*(4), 1193-1200.
- Mahoney, J. L. (2000). School extracurricular activity participation as a moderator in the development of antisocial patterns. *Child Development, 71*(2), 502-516.
- McNeil, L. M., Coppola, E., & Radigan, J. (2008). Avoidable losses: High-stakes accountability and the dropout crisis. *Education Policy Analysis Archives, 16*(3).
- Moretti, E. (2007). Crime and the costs of criminal justice. In C. R. Belfield & H. M. Levin (Eds.), *The price we pay* (pp. 142-159). Washington, DC: Brookings Institution Press.
- Muenning, P. (2007). Consequences in health status and costs. In C. R. Belfield & H. M. Levin (Eds.), *The price we pay* (pp. 125-141). Washington, DC: Brookings Institution Press.
- Muthén, B. O. (2004). Latent variable analysis: Growth mixture modeling and related techniques for longitudinal data. In D. Kaplan (Ed.), *The Sage handbook of quantitative methodology for the social sciences* (pp. 345-370). Thousand Oaks, CA: Sage Publications.
- Muthén, B. O., Brown, C. H., Masyn, K., Jo, B., Khoo, S.-T., Yang, C.-C., . . . Liao, J. (2002). General growth mixture modeling for randomized preventive interventions. *Biostatistics, 3*(4), 459-475.
- Muthén, B. O., & Muthén, L. K. (2000). Integrating person-centered and variable-centered analyses: Growth mixture modeling with latent trajectory classes. *Alcoholism: Clinical & Experimental Research, 24*(6), 882-891.
- Muthén, L. K., & Muthén, B. O. (2007). *Mplus: Statistical analysis with latent variables user's guide* (5th ed.). Los Angeles, CA: Muthén & Muthén.
- Nagin, D. S. (2005). *Group-based modeling of development*. Cambridge, MA: Harvard University Press.
- NCES. (n.d.). Education longitudinal study of 2002 (ELS:2002), from <http://www.nces.ed.gov/surveys/els2002/>
- 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: A Multidisciplinary Journal, 14*(4), 535-569.
- Pallas, A. M. (1989). Conceptual and measurement issues in the study of school dropouts. In K. Nambodiri & R. G. Corwin (Eds.), *Research in the sociology of education and socialization* (Vol. 8, pp. 87-116). Greenwich, CT: JAI Press Inc.
- Rasinski, K. A., Ingels, S. J., Rock, D. A., Pollack, J. M., & Wu, S.-C. (1993). America's high school sophomores: A ten year comparison. Washington, DC: National Center for Education Statistics, Office of Educational Research and Improvement, U.S. Department of Education.
- Raudenbush, S. W., & Bryk, A. S. (2002). *Hierarchical linear models: Applications and data analysis methods* (2nd ed.). Thousand Oaks: Sage.
- Ream, R. K., & Rumberger, R. W. (2008). Student engagement, peer social capital, and school dropout among Mexican American and Non-Latino White students. *Sociology of Education, 81*(2), 109-139.
- Roderick, M. (1994). Grade retention and school dropout: Investigating the association. *American Educational Research Journal, 31*(4), 729-759.
- Rouse, C. E. (2007). Consequences for the labor market. In C. R. Belfield & H. M. Levin (Eds.), *The price we pay* (pp. 99-124). Washington, DC: Brookings Institution Press.
- Rumberger, R. W. (1987). High school dropouts: A review of issues and evidence. *Review of Educational Research, 57*(2), 101-121.
- Rumberger, R. W. (1995). Dropping out of middle school: A multilevel analysis of students and schools. *American Educational Research Journal, 32*(3), 583-625.
- Rumberger, R. W. (2004). Why students drop out of school. In G. Orfield (Ed.), *Dropouts in America: Confronting the graduation rate crisis* (pp. 131-156). Cambridge, MA: Harvard Education Press.
- Rumberger, R. W., & Palardy, G. J. (2005). Test scores, dropout rates, and transfer rates as alternative indicators of high

- school performance. *American Educational Research Journal*, 42(1), 3-42.
- Swanson, C. B. (2004). Sketching a portrait of public high school graduation: Who graduates? Who doesn't? In G. Orfield (Ed.), *Dropouts in America: Confronting the graduation rate crisis* (pp. 13-40). Cambridge, MA: Harvard Education Press.
- Tyler, J. H. (2003). Economic benefits of the GED: Lessons from recent research. *Review of Educational Research*, 73(3), 369-403.
- Vermunt, J. K., Tran, B., & Magidson, J. (2008). Latent class models in longitudinal research. In S. Menard (Ed.), *Handbook of longitudinal research: Design, measurement and analysis* (pp. 373-385). Burlington, MA: Elsevier.
- Waldfogel, J., Garfinkel, I., & Kelly, B. (2007). Welfare and the costs of public assistance. In C. R. Belfield & H. M. Levin (Eds.), *The price we pay* (pp. 160-174). Washington, DC: Brookings Institution Press.
- Wang, M., & Bodner, T. E. (2007). Growth mixture modeling: Identifying and predicting unobserved subpopulations with longitudinal data. *Organizational Research Methods*, 10(4), 635-656.
- Wehlage, G. G., Rutter, R. A., Smith, G. A., Lesko, N., & Fernandez, R. R. (1989). *Reducing the risk: Schools as communities of support*. Philadelphia, PA: The Falmer Press.
- Zvoch, K. (2006). Freshman year dropouts: Interactions between student and school characteristics and student dropout status. *Journal of Education for Students Placed at Risk*, 11(1), 97-117.

Appendix 1: Variable Descriptives

Variable name	Mean	SD	Min.	Max.	Description (ELS variable)
Dropout	0.09	0.29	0.00	1.00	Dropped out at any time (F2EVERDO)
GPA9S1	2.90	0.82	0.00	4.00	Non-cumulative grade point average grade 9 semester 1
GPA9S2	2.85	0.86	0.00	4.00	Non-cumulative grade point average grade 9 semester 2
GPA10S1	2.82	0.89	0.00	4.00	Non-cumulative grade point average grade 10 semester 1
Female	0.51	0.50	0	1	BYSEX=2 (Male = 0, Female =1)
African American	0.14	0.34	0	1	BYRACE=2
Asian	0.11	0.31	0	1	BYRACE=3
Hispanic	0.07	0.25	0	1	Any Hispanic subgroup composite (BYHISPAN)
Non-traditional Family	0.22	0.41	0	1	Whether less than two parents or guardians in the home (BYFCOMP)
SES	0.13	0.62	-0.99	1.80	Socio-economic status composite (BYSES1)
Extracurricular	5.03	5.88	0	21	Hours per week spent on extracurricular activities (BYS42)
Retained	0.11	0.31	0	1	10 th grader ever held back a grade self report (BYP46)
Negative Behavior	0.58	0.48	0	4	Negative behavior composite
% Free lunch	19.14	15.60	0	95.14	Percent free lunch 2000/01 (CP01FLUN)
Urban	0.20	0.40	0	1	BYURBAN=1
Rural	0.28	0.45	0	1	BYURBAN=3
Student-teacher ratio	0.00	4.14	-10.39	61.91	Student-teacher ratio, grand mean centered (CP01STRO)
Academic Press	3.01	0.79	0	4	Teachers press students to achieve (0-4) as reported by the principal (BYA51B). 0=not accurate, 4=very accurate.
Small school	0.22	0.42	0	1	Enrollment < 600 (CP01STEN)
Large school	0.25	0.43	0	1	Enrollment 1201-1800 (CP01STEN)
Extra-large school	0.23	0.42	0	1	Enrollment >1800 (CP01STEN)

Appendix 2: Principal Component and Factor Loading Description for Negative Behavior Variable.

Item label	Item description	Loading
Negative Behavior		
BYS24A	How many times late for school	0.598
BYS24B	How many times cut/skip class	0.674
BYS24C	How many times absent from school	0.509
BYS24D	How many times got in trouble	0.703
BYS24E	How many times put on in-school suspension	0.727
BYS24F	How many times suspended/put on probation	0.708
BYS24G	How many times transferred for disciplinary reasons	0.461
Cronbach's alpha		0.701

Note: All variables were student self reported and were coded as: 0=*never*, 1=*1-2 times*, 2=*3-6 times*, 3=*7-9 times*, 4=*10 or more times*.

Appendix 3: MPLUS Code:

```

TITLE:      Growth Mixture Model for Dropout Identification

DATA:      FILE = GMM 01.dat;

VARIABLE:  NAMES = BYSEX aframer asian
            hispanic nontradf BYSES1 urban rural
            extracur retain freelnch smllenrl lgenrl xlenrl
            stratio acapress negbehav
            GPA9S1 GPA9S2 GPA10S1 F2EVERDO;
MISSING = ALL (999);
IDVARIABLE = ID;
USEVARIABLES
            BYSEX aframer asian
            hispanic nontradf BYSES1 urban rural
            extracur retain freelnch smllenrl lgenrl xlenrl
            stratio acapress negbehav
            GPA9S1 GPA9S2 GPA10S1 F2EVERDO;

            CENTERING GRANDMEAN(stratio);
            CATEGORICAL = F2EVERDO;
            CLASSES      = c(4);

ANALYSIS:  ESTIMATOR      = MLR;
            TYPE          = MIXTURE MISSING;
            PROCESSORS    = 4;
            MITERATION    = 5000;
            STARTS        = 100 10;
            STITERATIONS  = 100;
            LRTBOOTSTRAP  = 100;

MODEL:     %OVERALL%
            intercepts slopes | GPA9S1@0 GPA9S2@1 GPA10S1@2;
            intercepts ON BYSEX aframer asian hispanic nontradf
                BYSES1 urban rural;
            slopes c ON BYSEX aframer asian hispanic nontradf
                BYSES1 urban rural extracur retain negbehav freelnch
                smllenrl lgenrl xlenrl stratio acapress;
            slopes WITH intercepts;

            %c#1%
            [intercepts* slopes*];
            intercepts ON BYSEX aframer asian hispanic nontradf
                BYSES1 urban rural;
            slopes ON BYSEX aframer asian hispanic nontradf
                BYSES1 urban rural extracur retain negbehav freelnch
                smllenrl lgenrl xlenrl stratio acapress;
            slopes WITH intercepts;

            %c#2%
            [intercepts* slopes*];
            intercepts ON BYSEX aframer asian hispanic nontradf
                BYSES1 urban rural;
            slopes ON BYSEX aframer asian hispanic nontradf
                BYSES1 urban rural extracur retain negbehav freelnch
                smllenrl lgenrl xlenrl stratio acapress;
            slopes WITH intercepts;

```

```
%c#3%
[intercepts* slopes*];
intercepts ON BYSEX aframer asian hispanic nontradf
    BYSES1 urban rural;
slopes ON BYSEX aframer asian hispanic nontradf
    BYSES1 urban rural extracur retain negbehav freelnch
    smllenrl lgenrl xlenrl stratio acapress;
slopes WITH intercepts;
```

```
%c#4%
[intercepts* slopes*];
intercepts ON BYSEX aframer asian hispanic nontradf
    BYSES1 urban rural;
slopes ON BYSEX aframer asian hispanic nontradf
    BYSES1 urban rural extracur retain negbehav freelnch
    smllenrl lgenrl xlenrl stratio acapress;
slopes WITH intercepts;
```

```
PLOT:      type is plot3;
           series is GPA9S1(1) GPA9S2(2) GPA10S1(3);
```

```
OUTPUT:    SAMPSTAT STANDARDIZED TECH1 TECH4 TECH7 TECH11 TECH12
           TECH14;
```

```
SAVEDATA:  SAVE=CPROBABILITIES;
           FILE IS CPROBSAV01.DAT;
           FORMAT IS FREE;
           ESTIMATES=MIXESTIMATES01.DAT;
```