Early Warning System Simulations of High School Dropout Propensity
Lauren M. Porter, Lauren P. Bailes
Department of Education Studies, The Ohio State University

Abstract

High school dropout rates are a concerning trend for educators, administrators, education policy makers, and communities. A promising effort to address dropout rates is an empirical approach known as Early Warning Systems (EWS). The EWS is a prediction model that determines an individual student’s risk of dropping out of high school. The prediction model is designed to alert educators and parents to the risk at levels at an early stage. Doing so allows for earlier interventions in order to shift the trajectories of students displaying risks for dropping out. To this point, EWS have been limited in their predictive power because they are typically single-level regression models with limited and often out-of-date sample information. In this study, we produce a simulated sample of students consistent with the existing covariance matrix of variables from Delaware's EWS report (Merola & Fernandez, 2010). This work integrates extant research on EWS models and the educational imperative to predict a student’s likelihood of school dropout. Dropout rates have been shown to significantly impact minority populations; however, many statewide dropout models do not incorporate racial identity as a predictor. In this study, we conduct a simulation model with three continuous predictors (attendance rate, number of disciplinary referrals, and English Language Arts (ELA) final score). In the second model, we then include student race along with the three continuous predictors in order to determine whether or not the model fit is significantly different with inclusion of student race. We believe the use of simulated modeling contributes to EWS development because it allows for testing characteristic or program effects in a way that is not directly tied to individual-level data. The experimental design of simulation studies allows for a level of control that results in a more precise extrapolation of population parameters from sample parameters.

Introduction

Despite rises in graduation rates nationwide, dropout remains a problem, especially for minority students. In fact, students of color account for a larger proportion of students who dropout of high school even accounting for increased graduation rates. Early Warning Systems (EWS), statistical models that predict student risk of dropout, hold promise as one method of identifying students so they can receive necessary interventions. However, these preventative assessments are recent (the earliest of such studies was published in 2007) with policy and practice implications still to be fully understood. EWS also come with political concerns related to variables included in the model and the risk of attendant signal and noise constraints, which may misidentify students as safe or at risk of dropout. A study of dropout predictors found, using
Relative Operating Characteristics (ROC) curves, that the majority of dropout flags have high precision but are not characterized by high accuracy, and the most accurate flags were course performance flags (grades) (Bowers, Sprott, & Taff, 2013). We believe the use of simulated modeling contributes to EWS development because it allows for testing characteristic or program effects in a way that is not directly tied to individual-level data yet allows for examination of different combinations of student characteristics as predictors of dropout risk. The central aim of this study is to simulate data for an EWS model with and without race in order to explore the differences in model fit and explanatory power between the two models.

**Literature Review**

We turn now to a review of extant literature. While concerns about dropout have existed for several generations, research on EWS is emergent. Because the work and the models are still so new, relatively few longitudinal applications are available for analysis. Instead, most of the existing literature focuses on model development and various models’ advantages in identifying at-risk students. It is within this body of work that we situate the current study. EWS have the potential to shift policies aimed at dropout prevention such that targeted interventions are implemented earlier, resulting in a more effective use of dropout prevention funds and, subsequently, increased graduation rates for all students. It stands to reason, then, that literature surrounding EWS has grown exponentially in its relatively short history, which is reviewed below.

When a student drops out of school, it is considered the culmination of process rather than a single event (Hammond, Linton, Smink, & Drew, 2007). Throughout this process, there are multiple ways to shift the trajectory of a student at risk of dropping out. Unfortunately, these efforts are often taken after the process has gained momentum to the point where the outcome –
and the lasting repercussions for a student’s life – cannot be reversed. High school dropout is more than just an education problem; it is a public policy problem. The detrimental effects of dropout extend not only the individual, but to his or her family, neighborhood, and society. Research indicates that school dropout is associated with lifelong and often irreparable consequences, including reduced earning power, geographical instability, ill health, and reduced well being (Sum, Khatiwada, McLaughlin, & Palma, 2009). One of the most promising big data approaches to the wicked problems of education may be the increasingly broad application of Early Warning Systems (EWS) for school dropout. Early Warning Systems (EWS) or Early Warning Indicators (EWI) are platforms that compile student-level information into a risk score or propensity of dropping out of high school.

High school dropout rates continue to be of great concern in education. The most recent federal data on high school graduation rates report that 19% of high school students do not graduate\(^1\). Table 1 explores graduation rates using NCES data by relevant student demographic characteristics. Disproportionately low graduation rates are seen among students with disabilities, Limited English Proficient students, American Indian, and Black students; all four of these groups have a graduation rate more than 10% below the national graduation rate.

\textit{Table 1: 2012-2013 4-year adjusted cohort graduation rate (ACGR), by race/ethnicity}

<table>
<thead>
<tr>
<th>State</th>
<th>Percent of students</th>
</tr>
</thead>
<tbody>
<tr>
<td>United States</td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>81.4</td>
</tr>
<tr>
<td>American Indian / Alaska Native</td>
<td>69.7</td>
</tr>
<tr>
<td>Asian / Pacific Islander</td>
<td>88.7</td>
</tr>
<tr>
<td>Hispanic</td>
<td>75.2</td>
</tr>
<tr>
<td>Black</td>
<td>70.7</td>
</tr>
<tr>
<td>White</td>
<td>86.6</td>
</tr>
<tr>
<td>Economically disadvantaged</td>
<td>73.3</td>
</tr>
<tr>
<td>Limited English proficiency</td>
<td>61.1</td>
</tr>
<tr>
<td>Students with disabilities</td>
<td>61.9</td>
</tr>
</tbody>
</table>

\(^1\) \text{http://nces.ed.gov/cd/tbl/ACGR_2010-11_to_2012-13.asp}
The earliest EWS research (Neild, Balfanz, & Herzog, 2007) indicates that among several cohorts of Philadelphia 6th graders, individuals demonstrating one of the four following risk factors had a three in four chance of dropping out of high school:

- Math grade of F
- English grade of F
- Attendance rate below 80%
- Behavior mark of ‘unsatisfactory’ in at least one subject

Eighth grade students in this sample had dropout risks significantly predicted by the attendance rate markers and course performance markers (English and Math) consistent with the 6th grade model. Attendance, behavior, and course performance (either test scores, number of earned credits, or course failure) are hallmarks of statewide EWS models (Allensworth & Easton, 2005; Consortium, 2011). This early work from Balfanz established a pattern of dropout prediction that has largely been adhered to since its emergence, and many now-public state models include some combination of student-level predictors that include the ABC’s of dropout: attendance, behavior, and course performance.

Typically, EWS transform a set of student data into a statistical model that predicts an individual student’s risk of dropping out of school. Many states and some large school districts have begun to quantify predictors of dropout, often attempting to pinpoint the individual characteristics that put a student at risk for dropping out of high school. In 2013, 31 states provided some early warning reports (Data for Action, 2013). In 2011, there were 18 states providing reports of this nature. This marked increase demonstrates growing demand for an empirically based approach to moving the needle on high school dropout rates. Longitudinal analyses using student level data from 6th grade academic records were able to predict 60% of
students who dropped out before graduation when this information was plugged into an EWS (Balfanz, Herzog, & Mac Iver, 2007). EWS represent one frontier of preventative assessment in the P-12 system and so it behooves researchers, policymakers, and practitioners to develop the best possible model for identifying risk and preventing dropout, even as student populations change quickly and dramatically across demography, geography, language, and cognitive need.

As the development and use of EWS proliferates across states and districts, evident variation emerges in terms of the kinds of models used to predict individual student dropout. At their most basic, EWS include a few demographic or academic performance indicators, such as attendance, grade point average (GPA), or number of disciplinary referrals in a linear regression. In response to the clustered effects that are seen in dropout trends, wherein large percentages of a state’s dropouts are concentrated in a small number of buildings (Balfanz & Legters, 2004), Arkansas uses a multi-level predictive model that includes student-level and school-level predictors (Balfanz & Byrnes, 2010) More complex models use a combination of social, academic, and demographic indicators to predict a student’s likelihood of dropping out. This type of model is restricted to states with a sufficient number of districts to serve as the Level-2 units of analysis. With only one multi-level EWS system publicly available, the warrant for use of single-level regression is stronger, as it can be utilized by states with any number of school districts. While this expanded model includes more predictors, all predictors are features of the student and not of the school or community. Additionally, key characteristics such as race or student disability status are not included in the dropout models.

Carl, Richardson, Cheng, Kim, and Meyer (2013) identify some limitations of binary on-track/off-track indicators conventionally used by previous dropout researchers (Allensworth & Easton, 2005, 2007). Such binary indicators are often not measured until a student’s second or
third year of high school, at which point many students have already begun the dropout process or dropped out altogether. A more useful measure of dropout risk is assessed at or before 9th grade. A second limitation is that binary measures do not distinguish between on-time graduation and graduation at all. That is, there is no way to tell how far off track a student is with regard to particular time-bound academic outcomes of interest. Finally, a binary indicator does not provide any information about a student’s risk of dropout at different times throughout his or her high school career, which reduces the sensitivity with which any student’s risk can be assessed over four years.

A better dropout assessment solution, posit Carl et al., would address each limitation: it would measure a student’s risk immediately upon (or before) entry into high school, provide risk assessment in relation to the goal of on-time graduation, and be sensitive to a student’s variable risk propensity over time. They propose a system of Total Quality Credits (TQC) as one way to measure student dropout risk. TQC is comprised of a linear relationship between credit attainment and final grades in academic classes based on a 0-4 scale. So, for example, a freshman who had taken four academic classes and received an A (4.0), 2 Bs (3.0), and 1 C (2.0) would have a TQC score of 12 (out of a possible 16). Not only does the TQC method have strong negative correlations with a series of disciplinary events, further confirming that dropout is a process and not an event, but it has the added benefit of mimicking the criteria by which students are judged for college admissions. First year high school TQC has significant predictive power when it comes to predicting not only on time graduation but also on time college enrollment (defined as matriculation by November 1 of a student’s graduation year). In fact, a TQC of 8 (or about a C average) seems to be the threshold at which student’s likelihood reaches approximately 70%. As students’ TQC increases above 8, their graduation likelihood seems to plateau – their
risk neither increases nor decreases substantially. While TQC offer some improvements on the most basic EWS models, Carl and colleagues acknowledge that TQC still has limitations. Specifically that it cannot account for engagement, personal, or social factors, which the researchers call stress factors, that have may contribute to a student’s risk of dropout. This study argues that, because race often correlates with other risk and stress factors such as poverty or mobility, the inclusion of race as a predictor would help identify more students of color who may be at risk for dropout.

Poverty is the key predictor of promoting power (a school’s ability to successfully see students through high school in four years) in a high school. A report by Balfanz and Legters (2004) presents a number of staggering statistics about the places where students are least likely to graduate and the kinds of students who are sentenced to schools where on-time graduation is, at best, a 50/50 proposition. Those students tend to identify as minorities: “Nearly half of our nation’s African-American students, nearly 40% of Latino students, and only 11% of white students attend high schools in which graduation is not the norm” (p. v). Despite the disproportional cost that minority students pay in terms of the likelihood of their graduation, race does not factor into existing EWS. This paper holds that, given the correlational nature of race, poverty, and dropout in this country, the addition of a race variable would improve the predictive power of EWS formulae.

This study holds that the addition of race to some Early Warning Systems would enhance a school’s power to predict dropout, primarily due to the unjust but undeniable associations among a student’s race and his or her likelihood to attend school where poverty and reduced promoting power are the norm. Given the disproportionate dropout trends amongst minority students, an unexpected gap in the literature exists when considering if the early warning systems
are as accurate for minority students as they are for white students. It stands to reason that this is related to the lack of race as a predictor variable in many models. The inclusion of race as a predictor variable in early warning system models may improve the accuracy of predictions, specifically for minority students who have a greater rate of dropping out of high school.

If Early Warning Systems are to be maximally effective for use in practice and in policymaking, researchers must continue to refine the models so that they can powerfully and accurately identify at-risk students. The goal of this study is to examine one possible refinement for EWS models nationwide – the use of race as a predictor of dropout.

Rationale and Research Questions

It is difficult to generalize from the results of EWS models outside of populations of interest due to the unique populations around which state-cased EWS systems are structured. As these populations change, the models must adapt to new characteristics in the student population being assessed for risk. However, Monte Carlo simulations allow researchers to better understand the variables that shift a student’s overall dropout risk, without the ethical or political implications of purposefully assigning students to specific profiles of risk in order to study the statistical implications. Using simulations allows for the reflection of natural variability of a modelled risk (Pouillot, Delignette-Muller, Kelly, & Denis, 2013). Monte Carlo simulations originated in Metropolis & Ulam’s 1949 article, with Markov Chain Monte Carlo simulations appearing soon after (Metropolis & Ulam, 1949). This technique uses computing power, and stated distributional properties, to convert random observations into deterministic results. Monte Carlo simulations are particularly useful in scenarios where model specifications are being tested with high-stakes outcome measures, such as course failure, lack of receiving intervention
treatment, or in the medical field, death. Simulations allow for the subjects assigned to certain outcome variables to test the model to be imaginary, and therefore avoid the ethical issues of intentionally assigning students to adverse outcome variables.

This study is designed to familiarize the audience with core elements of a dropout prediction model. This study will explore variability in models through a simulation study. Use of simulated data allows for the exploration of race as an additional predictor of individual dropout risk.

This paper is designed to answer the following research question:

- How does the model fit of an early warning system model change with the addition of race as a predictor of high school dropout risk?

**Primary Analytic Method**

Variables for this study include attendance rate, number of disciplinary offenses, English Language Arts (ELA) final grade, and race. The correlation matrix in Table 3 from Delaware’s Early Warning System (Merola & Fernandez, 2010) was used to simulate 2,000 cases using the publicly available covariance matrix for attendance rate, number of disciplinary offenses, and ELA final grade in the Delaware Early Warning System report. Data were simulated using R 3.2.2. Simulated cases were then randomly assigned race designations based on the breakdown of race of enrolled students in the Delaware report. Within the racial categories, cases were assigned to dropout or non-dropout categories randomly using the percentages of Delaware students within each racial category that dropped out of high school in the 2012-2013 school year (Table 2).

*Table 2: Delaware Enrollment and Dropout Rates, By Race*

<table>
<thead>
<tr>
<th>Race</th>
<th>Enrollment Total</th>
<th>Percent of Enrollment Total</th>
<th>Annual Dropout Rate</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Table 3: Covariance Matrix

<table>
<thead>
<tr>
<th></th>
<th>Attendance Rate</th>
<th>Number of Suspensions</th>
<th>Number of Offenses</th>
<th>Math Final Grade</th>
<th>ELA Final Grade</th>
</tr>
</thead>
<tbody>
<tr>
<td>Attendance Rate</td>
<td>1</td>
<td>-0.24</td>
<td>-0.15</td>
<td>0.35</td>
<td>0.36</td>
</tr>
<tr>
<td>Number of Suspensions</td>
<td>-0.24</td>
<td>1</td>
<td>0.39</td>
<td>-0.26</td>
<td>-0.26</td>
</tr>
<tr>
<td>Number of Offenses</td>
<td>-0.15</td>
<td>0.39</td>
<td>1</td>
<td>-0.15</td>
<td>-0.15</td>
</tr>
<tr>
<td>Math Final Grade</td>
<td>0.35</td>
<td>-0.26</td>
<td>-0.15</td>
<td>1</td>
<td>0.59</td>
</tr>
<tr>
<td>ELA Final Grade</td>
<td>0.36</td>
<td>-0.26</td>
<td>-0.15</td>
<td>0.59</td>
<td>1</td>
</tr>
</tbody>
</table>

Table 3’s Covariance matrix is based on individual-level incident counts for numbers of suspensions and offenses, individual-level attendance rate (percentage formed from subtracting the number of days absent from total days), and standardized final grades for Math and ELA subject areas.

A critical consideration when understanding the simulated model built here is that it is not intended as a replication of Delaware’s own EWS model. Delaware’s published EWS model (Equation 1) does not include number of disciplinary offenses. Instead this model is intended to broadly represent common forms of dropout prediction models nationwide, many of which include some combination of student-level predictors such as attendance, behavior, and course performance.

Equation 1: \( \text{Probability} = \beta_0 + \beta_1 * \text{attendance\_rate} + \beta_2 * \text{repeating\_gradesi} + \beta_3 * \text{reading\_final\_grade\_2009i} \)
The basic model for predicting dropout propensity includes attendance rate, number of disciplinary offenses and ELA Final Score as measures of the traditional attendance, behavior, and course performance predictors often seen in EWS models. The expanded model will include all predictors from the basic model, with the addition of race as a predictor of dropout propensity.

Single level regression models were then created for the base model as well as the model with the additional predictor of race. These models are to be examined for similarities and differences relative to model fit and amount of variance in dropout propensity explained by both the basic and expanded models.

**Results**

Table 4 provides evidence that both the basic and expanded models are suitable fits for the data, using the .05 criteria set forth by the Hosmer and Lemeshow test.

<table>
<thead>
<tr>
<th>Step</th>
<th>Chi-square</th>
<th>df</th>
<th>Sig.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Basic</td>
<td>5.651</td>
<td>8</td>
<td>.686</td>
</tr>
<tr>
<td>Expanded</td>
<td>9/166</td>
<td>8</td>
<td>.329</td>
</tr>
</tbody>
</table>

Table 5 demonstrates that when comparing the 2 different models, the expanded model is a slightly better fit. The -2 log likelihood (-2LL) decreases by less than 10 points with the addition of race. The Nagelkerke and Cox & Snell $R^2$ statistics are not appropriate for use in cases with continuous predictors; therefore, with the -2LL as the only comparative statistic for model fit, there is not enough evidence considering only the model summary to determine if the addition of race as a predictor results in an improved model.

<table>
<thead>
<tr>
<th>Table 5: Model Summary</th>
</tr>
</thead>
<tbody>
<tr>
<td>Step</td>
</tr>
<tr>
<td>Chi-square</td>
</tr>
<tr>
<td>df</td>
</tr>
<tr>
<td>Sig.</td>
</tr>
<tr>
<td>Basic</td>
</tr>
<tr>
<td>Expanded</td>
</tr>
</tbody>
</table>
Table 6 shows that when considering the typical attendance, behavior, and course performance predictors, only the measure of behavior (number of offenses) is a statistically significant predictor of dropout propensity. For every increased behavioral offense, a student can be expected to have an increased dropout risk of .760 units, controlling for all other predictors.

Within Table 6, there is evidence that the significance of Number of Offenses within the Basic model was in part due to its capturing the effects of race. Once Race was included as a predictor, Number of Offenses is no longer a statistically significant predictor of dropout propensity, but White students do have a statistically higher risk of dropout risk, and Black
students can be expected to have a risk 2.3 times higher than that of White students, controlling for all other predictors in the model.

Discussion

As the proliferation of EWSs occurs throughout states and districts, it is increasingly important to understand how these predictive tools vary depending on student characteristics and policy contexts. There has not yet been a comprehensive and quantitative evaluation of EWS across regions and school systems, in part because the models are not intended to generalize outside of the original population intended by the model. The basic and expanded models in this study share no variables that significantly predict dropout propensity, however there are variables that are non-significant across both models (attendance rate, ELA Grade) that would have been expected to be significant, considering only the literature on Early Warning System models. Our study highlights the differing potency of variables in predicting dropout in different populations, demonstrating the benefits of employing simulation in developing statewide models. Simulating the effects of variable inclusion in the model allow for flexibility in adjusting the covariance matrix and understanding the results of changing relationships between predictor variables.

State and local governments have experienced mounting pressure to identify not only the risk factors associated with dropout but to produce the data by which at-risk students may be identified (Balfanz & Legters, 2004). High school completion is seen to be a valuable measure of student achievement for several reasons: it is an unambiguous measure of student attainment, it is relatively simple to compare graduation rates of similar populations, and late graduates cost their locales more in terms of intervention and remediation (Knowles, 2014). Conversely, on-time high school graduation is related to a constellation of benefits for individuals and
communities, ranging from psychological well-being, to greater earning power, and increased financial stability. State, local, and school-level administrators are being held accountable for discerning which educational programs and initiatives increase student persistence and high school completion. Policy and governance structures in some states reflect this pressure. In Delaware, for example, the state’s Department of Education formed the Delaware Promise Dropout Prevention Subcommittee. The purpose of this subcommittee was to deploy America’s Promise grant funds in pursuit of a functional and accessible dropout early warning system.

As of 2013, 31 states provide some early warning reports (Data for Action, 2013). In 2011, there were 18 states providing reports of this nature. However, there is still great need for early warning systems that provide information about risk factors to a number of stakeholders that may include students, parents, guidance counselors, teachers, administrators, and policymakers. Also, despite the progress states have made in generating risk assessments, there needs to be greater attention paid to programmatic failures. That is, Early Warning Systems need to identify whole schools, districts, or regions that fail to support students’ high school completion.

Models are developed for the specific contexts of states, therefore must be adaptable for the changing demographics and policies within a state. Lack of fit for sample data is reflective of the fact that state-defined models are by necessity not generalizable to others states, which have distinct policy environments and student bodies. Differences in state policies related to time in school, school attendance, graduation requirements, content standards, teacher license and assessment are detailed in The Council of Chief State School Officers 2009 report *Key State Education Policies on PK-12 Education: 2008*. Some states have local policy, while some have statewide guidance on the topics, but there is a great deal of variation on policies directly related
to graduation requirements as well as student attendance, an often-used EWS indicator (CCSSO, 2008).

A snapshot of the changing student demographics can be seen in NCES’ *Condition of Education* (Kena et al., 2014). Showing the national increases through the lens of individual state enrollment increases and decreases, as well as the demographic characteristics of students in those states, it becomes clear that models will necessarily need to be flexible and adaptable in order to accommodate enrollment changes. For example, national public school enrollment increased by 2.3 million from 2000-2001 and 2011-12 school years; this overall includes steeply increasing states such as Nevada with a 29% increase in enrollment steeply decreasing states such as Vermont (12% decrease). Models that are designed for a state in 2001 may not fit after ten years of state policy changes compounded with enrollment changes within the state. Models must be adaptable, specific to changes in student demographics as well as school level variables that change over time.

**Limitations**

The consideration of characteristics by student race is limited by small cell sizes in certain districts. Disciplinary action data was not available disaggregated by race; this data was only publicly available at the building or district level.

Some risk factors are building level, and those are not addressed in this piece. Future research in this area would do well to explore the risks that exist associated with the school, separated from the risks associated with individual risk factors, as outlined in the Knowles and Carl articles as the warrant for the stress test suggested.
References


