Identifying Early Indicators for College Readiness

Hansheng Chen

Agency Fellow

Los Angeles Unified School District

Introduction

My working experience with Los Angeles Unified School District (LAUSD) before I joined Strategic Data Project (SDP) at Harvard University can be classified into two different categories: research and data warehouse. In the early years, I worked with LAUSD's Program Evaluation and Research Branch as a traditional research analyst. I played a role following typical research procedures such as developing hypotheses and research designs, collecting data, conducting analyses, and finding evidence to support or reject the hypothesis. Later, I worked primarily on the development of a then newly established Decision Support System (DSS), an Oracle based central data warehouse in the district.

My work included the development of an Early Alert System that monitors students' performance across the district using multiple measures in both academic and behavioral areas. The system consists of two sub-systems: 1) an At-Risk System that identifies at-risk students using a carefully defined "at-risk zone," and 2) an Early Warning System that triggers an "early warning" by observing the relative changes in student performance.

I also designed an A-G Course On-Track System, which was put in use in 2007 to diagnose whether students are on-track to fulfill college requirements. A-G courses are seven categories of high school courses certified by California public universities as necessary for college admission. The system defines the logic of course sequence and projects the general pattern the student should follow to keep on track to complete the college admission pre-requisites grade by grade.

Other work I did included building and applying a logistic regression model to predict dropouts. The model was built on a sample of 174,885 students and applied to 202,437 students with an overall accuracy of 70.1%. I also worked on data mining model estimating the passing rates on California High School Exit Exam (CAHSEE). The model retrieved 246 attributes as predictors. The model was run directly from the data warehouse and generated an overall 81.2% accuracy. Building a data mining model was an effort to borrow the non-traditional research methodology that had been used successfully in other businesses such as financial institutions and health care, and apply it to public education.

SDP Influence on High School Exit Exam Analysis

Since I joined the SDP fellowship, besides the ten regular workshops I attended in the past two years, there are several things that I feel were particularly helpful. One is the assignment by our reading club to read the book *Methods Matter* by Richard Murnane and John Willett. I was impressed by the methods they described in the book, such as the natural experiments, difference in difference, and instrumental-variables estimation (IVE). The book was both challenging and readable. Many statistical concepts in the book, especially those addressed in the footnotes, are explored deeply but expressed clearly. SDP Faculty Advisor Martin West facilitated a few discussions of chapters of the book. I had very interesting discussions with him, via phone and email, on validity, type I error, sample size, effect size, statistical power, etc. The communication was thought-provoking. I also had a chance, on several occasions, to consult with Meredith Phillips, our faculty advisor, on covariance and ordinal regression, the latter of which was specific to one of my projects. In addition, two other books we read were beneficial: *Measuring Up* and *Leading for Equity*. Reading *Measuring Up* was a quite enjoyable experience under the current atmosphere dominated by value-added models. Considering Koretz's "politically incorrect" truth, which if not explicitly was at least implicitly addressed in his book, was pretty amusing. Although less philosophical compared to *Measuring Up*, *Leading for Equity* is more practical. As a matter of fact, I was able to apply some of the ideas in the book into one of my projects.

Borrowing the idea of the "seven keys to college readiness" in *Leading for Equity*, I tried to find early indicators of college readiness in our district. Taking cohort 2011 as an example, I found 5^{th} grade advanced reading and completion of Algebra 1 by 8^{th} grade are indeed strong indicators for college readiness in our district. Cohort 2011 was classified into four categories: 1) dropout, 2) 4-year on track graduation, 3) on track graduation plus meeting A-G requirement, and 4) on track graduation plus meeting A-G requirement plus SAT scoring at 1650 or higher. The four categories represent four groups of students: dropout, graduate, graduate meeting 4-year college requirement, and graduate meeting selective college requirement.

When cohort 2011 student data was matched with reading achievement data in 2004 as measured by California Standards Test (CST) English Language Arts performance level,[1] I found that while 7% of those who were reading proficiently later dropped out, 17% of those who were not reading proficiently dropped out. The difference became even larger for the other three categories of students: 76% of the proficient reading group and 51% of non-proficient reading group graduated; 37% of the proficient reading group and 12% of the non-proficient reading group graduated and met A-G; 12% of the proficient reading group and 0% of the non-proficient reading group graduated, met A-G, and scored at 1650 or higher for the SAT. The relationship between 5^{th} grade reading achievement level and later graduation and college readiness status is statistically significant at 0.000 level for all four categories.

[1] Reading proficient is defined as CST ELA performance level at proficient or higher, non-proficient as basic or below.

A similar robust relationship was found when the 8th grade Algebra I data were matched. Eight percent of those who completed Algebra I scoring C or better by 8th grade would drop out and 18% of those who were not able to complete Algebra I scoring C or better dropped out. The difference also became larger for the remaining three categories of students: 78% of the Algebra I group and 47% of the non-Algebra I group graduated; 42% of Algebra I group and 8% of the non-Algebra I group graduated and met A-G; 10% of Algebra I group and 0% of non-Algebra I group graduated, met A-G, and scored 1650 or higher for SAT. Again, the relationship between 8th grade Algebra I performance and later graduation college readiness status is statistically significant at 0.000 level for all four categories.

Another effort to explore early indicators for on track graduation was to link the existing Early Alert System in our data warehouse with graduation rate. When cohort 2011 graduation and dropout data were linked with 9th grade student risk factors in the data warehouse, a pretty strong relationship can be identified. While 93% of students who had 0 risk factors and 73% of students who had 1 to 4 risk factors at 9th grade later graduated in 4 years, only 31% of those who had more than 5 risk factors at 9th grade graduated in 4 years. A similar relationship was found between the risk factors and dropout rate: 5% of 0 risk factor students, 16% of 1 to 4 risk factor students, and 29% of more than 5 risk factor students dropped out during their high school years. A logistic regression model was built for the risk factor and graduation rate with an overall accuracy of 72%. Another logistic regression model for the risk factor and dropout rate had 66% accuracy. I presented the findings at one of SDP's workshops.

Tenth grade test results of CAHSEE directly affect graduation rate. To raise graduation rates, the district implemented a CAHSEE diagnostic test for 9th graders a few years ago. I conducted a series of analyses to examine the effectiveness of the diagnostic test. For example, there were 28,737 students who took the math diagnostic test at 9^{th} grade in 2010 and CAHSEE math the following year. Among those students, 63% (13,148) who failed the math diagnostic test in 2010 passed CAHSEE math in 2011. The diagnostic test was a huge help in preparing students for passing CAHSEE. However, the “progress” was too dramatic to be true. Was the huge “progress” a result of effective intervention following the diagnostic test? Or was it primarily caused by the “false positive” prediction (i.e. predicting more students failed than actually failed) of the diagnostic test?

The “progress” was dramatically reduced when we matched the test data with the intervention data. Among those who failed the math diagnostic test, while 74% of students passed CAHSEE math with intervention, there were also 62% of students who passed CAHSEE math without intervention. The 62% of students who were diagnosed as failed but actually passed CAHSEE math without intervention demonstrated the “false positive” bias of the diagnostic test. However, the existing difference of 12 percentage points between intervention and non-intervention groups could still be somehow attributed to the diagnostic test, assuming the intervened students would not get interventions without the diagnostic test.

The difference of 12 percentage points was further reduced to 8.5 percentage points if we control the student prior CST score. While 93% of students with high CST scores[2] passed CAHSEE with intervention, 90% of students with high CST scores also passed CAHSEE without intervention. The difference amongst the low CST score students, however, was much larger. While 68% of students with low CST scores passed CAHSEE with intervention, only 54% of students with low CST scores passed CAHSEE without intervention. Among the high

[2] High CST score is defined as performance level at basic or above, and low CST score as below basic or far below basic.

CST score students, the difference between intervention and non-intervention groups was only 3 percentage points, much smaller than the 12 percentage point difference before CST was controlled. Amongst the low CST score students, the difference between intervention and non-intervention groups was 14 percentage points, even larger than the original 12 percentage points. This finding raised a question: Is CST a better indicator for intervention? Or, more generally, is CST a better indicator than CAHSEE diagnostic test for predicting students' CAHSEE results? If this is the case, why should our district spend additional resources to implement the CAHSEE diagnostic test?

To address these questions, a comparison of the predictive power on CAHSEE results between the diagnostic test and CST was conducted. In both English and math, CST predicted CAHSEE results better than the diagnostic test. In the case of math, the diagnostic test generated 50% overall accuracy and CST had 72% overall accuracy. We know that the diagnostic test has a "false positive" bias, maybe on purpose in its design, so that more students who might potentially fail CAHSEE can be provided interventions in order to ensure the highest rate of CSHSSE success. To have a fair comparison, the CST prediction was adjusted to produce a lower "false negative" number. For example, in the first math comparison, although CST prediction had a higher overall accuracy, there were 722 students with high CST who failed CAHSEE. This number is much higher than the 152 students who passed the diagnostic test but failed the actual CAHSEE. A fair comparison in this case is to lower the 722 students to a point that is equal to or lower than the 152 students, and then compare the overall accuracy. When the 722 students were eventually lowered to 147 students, CST prediction still has an overall accuracy one percentage point higher than the diagnostic prediction. We then concluded that CST predicted CAHSEE results better than the diagnostic test.

Several regression models were built to estimate the CAHSEE results for at-risk students based on their prior CST scores. A new report will be created in the data warehouse based on these estimations, and it will be refreshed whenever the 9th grade CST data is available to help prepare students for the CAHSEE test. Since the prediction based on CST score is more cost effective, the CAHSEE diagnostic test will be discontinued.