education analytics

Equity-Aligned Analytics to Support Integrated Early Warning and School Accountability Systems

Robert H. Meyer, Rinor Jahiu, Anthony T. Milanowski, and Jordan B. Owen

DOI.ORG/10.62137/PBOG7141

110 E. Main St.
Suite 1000
Madison, WI 53703

CONTENTS

*This report is part of a series of research on innovative assessment approaches and alternative accountability models coordinated under The **K12 Research for Equity Hub**. The Hub is managed by **EduDream** and funded by the Bill & Melinda Gates Foundation and the Walton Family Foundation. No personnel from the Bill & Melinda Gates Foundation, nor the Walton Family Foundation participated in the creation of Hub research. The findings and conclusions contained in this report are those of the authors and do not necessarily reflect the positions and/or policies of the Bill & Melinda Gates Foundation or the Walton Family Foundation.*

INTRODUCTION

This report describes an approach to school accountability systems that addresses equity concerns by combining prospective measures of student readiness such as those found in early warning systems with a refined set of school- and district-based retrospective accountability measures based on improving student readiness. The major innovation in what we call the Equity-Aligned Analytics System (EAAS) is to provide validated indicators of projected readiness for high school graduation to students, parents, and school staff so they can take action to improve student outcomes, rather than awaiting results from conventional post-hoc accountability measures. These student-level readiness indicators are then rolled up to the school and district levels, and the improvements in readiness between years become part of the school and district accountability system.

Theory of Action

The basic theory of action for EAAS is that providing prospective analytics projecting important medium-term outcomes for individual students will catalyze educators, families, students, and other key stakeholders to focus efforts on specific short-range outcomes (i.e. Attendance, Course Enrollment, Test Scores and Grades) that lead to improvements in the medium-range outcomes such as high-school graduation. In parallel, the projections are used to create retrospective school and district accountability measures that show whether these efforts have succeeded in improving students' chances of experiencing positive outcomes.

By using the changes in projections from the end of one year to the next as the foundation of a school accountability system, this approach provides an incentive for schools to improve projected readiness as well as tools (detailed projections for individual students) for understanding how to improve readiness for individual students. The alignment of the prospective and retrospective metrics means that the same actions school staff and other stakeholders take based on indicators of student readiness also directly contribute to improving their school's overall accountability standing. The prospective metrics illustrate potential paths through which retrospective school and district-focused accountability measures can be improved. This approach is equity-oriented because it focuses attention on what each individual student needs to succeed and facilitates explicit assessment of schools' and districts' relative success with underserved and marginalized groups.

The EAAS expands the utility and equity orientation of a typical school accountability system in several substantial ways.

1. EAAS demonstrates the feasibility and appropriateness of expanding accountability metrics beyond test scores and similar aggregate indicators. The expanded set would include a broader set of student indicators, such as attendance, enrollment in challenging courses, and course grades.

2. EAAS integrates early warning and school accountability system features to create a cycle of continual measurement of student progress and support, much like a multi-tiered system of supports. It provides a roadmap to allow educators, students, and parents to maximize opportunities for improvement and to open pathways to future life success. It allows students to compensate for weaknesses in one indicator if they excel in other directions.

3. EAAS makes explicit within- and between-school differences in student outcomes and progress for different demographic groups. Comparative readiness and progress estimates for demographic groups of interest (e.g., economically disadvantaged versus not economically disadvantaged) help schools and the community know whether all students are learning to their full potential within the school environment.

To illustrate how this approach would be implemented and demonstrate its feasibility, EA worked with a large metropolitan district to develop early warning indicators for high school graduation, created student and school-level readiness and progress metrics, and created illustrative reports that can be used as part of ongoing early warning and school accountability systems. The reports focus on student outcomes and readiness in 8th and 9th grades. When fully developed, EAAS would be implemented in all grades.

Report Roadmap

The rest of this report consists of four sections. Section I briefly reviews the basics of early warning and accountability systems to show why these two tools could benefit from a synergistic combination. Section II provides an overview of the underlying methodology of EAAS and how it differs from existing systems. Section III illustrates our approach using work we have done for a large US school district. While the example is based on a specific district, implementing a similar system in other districts or states using different measures would follow the same principles. We describe this work in detail to illustrate what a system would look like and as a vehicle to discuss the issues that need to be addressed in implementing our approach. Since the district has not yet put the system to use, we cannot share any evidence as to its efficacy. Section IV discusses what we have learned in developing our approach to date, its potential limitations, and how we plan to expand the work.

I. REVIEW OF ACCOUNTABILITY AND EARLY WARNING SYSTEMS

Typically, early warning systems and school accountability are seen as separate tools within American education. Accountability systems typically produce school-wide scores on several dimensions that reflect last year's school performance. In contrast, early warning systems are focused on individual students, using past achievement and other indicators to predict future outcomes, often including high school graduation. Each has evolved separately, and few interconnections exist.

Early Warning Systems

Early warning systems are a form of **prospective analytics**, or forward-looking indicators. They give students, parents, educators, and policy makers insight into students' risks and opportunities so they may make informed decisions. This type of analytics prompts inquiry into the question, "What could happen?" and enables planning and execution of strategies to support students and stakeholders. As such, they must be provided at the beginning of the school year and, more generally, prior to decision events during the year, including, for example:

- Student selection of future academic, career, and workforce pathways and courses.
- Budgeting and planning to manage supplemental tutoring, after school, and summer school enrollment and instruction.

Our review of the literature suggests that early warning systems are most used to identify middle or early high school students who are at risk of not graduating from high school. They use factors like attendance, behavior, and course grades to predict whether a student is on-track to graduate. Though there are other important milestones or gateways that can be the focus of EWSs (e.g., third grade reading proficiency), high school graduation is especially important since it is typically the gateway to further educational or work opportunities that lead to greater adult wellbeing. While the research base is not extensive, it does suggest that early warning systems can accurately identify students at risk (McMahon & Sembiante, 2020; Pierson, Frazelle & Mazzeo, 2020; Wentwoth & Nagaoka, 2020; Carl et al, 2013; Johnson & Semelroth, 2010).

The theory of action underlying EWSs is that if students at risk (for example, of not graduating) can be identified before the fact, and that information is provided to school staff, they will take actions to support these students. In turn, these actions will lead to improvements for the identified students on the EWS predictors, which will lead to better outcomes for them.

One thing missing in this theory of action is the role of parents and students themselves. If we assume they have agency in improving their life chances, then the EWS should also provide them with information they can use to make decisions and take actions to improve the chances of more positive outcomes, including advocating for needed support.

Another missing piece is the need to tie EWS into other school practices and systems to encourage use (Mac Iver et al, 2019; Fox & Balfanz, 2020). EWS's implemented appear to have little or no explicit links with school accountability systems, even though better outcomes for identified students could improve schools' standings in the state/district accountability systems. Early warning system results are not made public and do not provide any incentives or consequences for use.

There is only limited evidence that EWS have impact on student outcomes (e.g., Perdomo et al, 2023; Corrin et al, 2017). The literature acknowledges that simply identifying at-risk students does not ensure action, and that supports such as professional development on data use, familiarity with and availability of interventions, and incentives for use are needed (Wentworth & Nakagoa, 2020; Fox & Balfenz, 2020; Frazelle & Nagel, 2015).

Accountability Systems

Accountability systems are intended to show how well schools have succeeded in meeting state or local performance goals, for providing parents with information on how well their children's schools are doing, and for motivating schools to improve performance. The theory of action underlying accountability systems is that setting performance goals, assessing students annually to see if goals are met, and providing incentives and consequences for meeting the goals (or not) will motivate schools (and districts) to pay attention to underserved students, and make changes to improve performance (Spurrier et al 2020b).

In contrast to early warning systems, accountability systems rely on **retrospective analytics** which enable review, reflection, diagnosis, and evaluation of student outcomes and progress in achieving these outcomes at all levels of the educational system. Retrospective analytics address the question, "What did happen?" and are intended to prompt continuous improvement of the enterprise of student learning.

Overall, it is not clear that current accountability systems have had the impact on educational equity that some of their original proponents intended (Torres, 2021; Spurrier et al, 2020; deBrey et al 2019; Harman et al, 2016). Among the prescriptions for improving the utility of accountability systems as tools for promoting equity are: including a broader set of measures (e.g., Cardichon & Darling-Hammond, 2017; Lee et al, 2019), focusing more explicitly on equity (Edley et al, 2019;), involving affected parties (e.g., families, educators) more in accountability system design (e.g., TNTP, 2016; Bush-Macenas et al, 2018), and focusing more on measures that contribute to success in later life (Cardichon & Darling-Hammond, 2017; Beach et al, 2015).

Embedding Prospective Analytics in Accountability

While there are several ways to address the shortcomings of both early warning and school accountability systems, the approach we exemplify in this report focuses on linking accountability with early warning systems, combining retrospective and prospective analytics, and focusing on medium term outcomes that are more closely related to life success than test scores alone. This combination is meant to help both systems become more effective in promoting equity in ways that impact students' futures.

To improve the impact of early warning systems, EAAS adds incentives by making progress in improving readiness a part of a school's accountability measure. It addresses shortcomings of accountability systems by expanding the set of student outcomes to those that predict high school graduation, an essential for future student success, and by providing schools (as well as students and parents) with a readiness profile that shows where to improve readiness. The readiness metrics, based on outcomes from the prior school year, form the basis for diagnosis, planning, and action to improve readiness over the new school year.

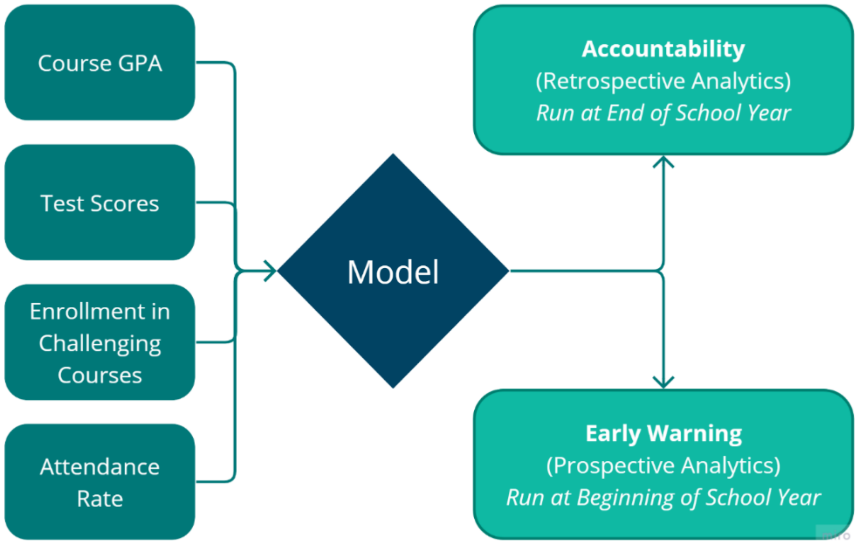
The combination can potentially advance equity by focusing attention on individual students' needs at a point when there is still time to improve, as well as reporting differences in readiness and progress in improving it for historically underserved groups. We illustrate examples of differences by economic disadvantage (ED) status and by race/ethnicity, but this could extend to other demographic groups in the future.

II. INTRODUCTION TO THE METHODOLOGY OF THE EQUITY-ALIGNED ANALYTICS SYSTEM

Constructing an Equity-Aligned Analytics System

In practice, prospective and retrospective analytics are based on the same fundamental process: student-level data is fed into a model that allows us to connect student outcomes in each year to projections (predictions) of future high school graduation status (or other medium and long-term outcomes). Prospective (forward-looking) analytics enable students, parents, and educators to take pro-active steps to support students and schools. Retrospective (backward-looking) analytics enable these agents to evaluate whether students and schools have made progress, measured in terms of contemporaneous high school outcomes and projected future outcomes. We refer to the system described here as the Equity-Aligned Analytics System (EAAS).

Figure 1. High-Level Summary of the System



First, we create the early warning part of the system: a statistical model that predicts future high school graduation status using historic student data, which is then used to project future high school graduation status based on information about current students. This involves two steps:

- First estimate a model of high school graduation status (dropout and non-graduate, regular (non-honors) diploma, or honors diploma) with four central predictors: attendance, enrollment in challenging courses, test scores, and course GPA; we then harvest the coefficients. These models are estimated using historic student longitudinal data.
- Then use these calibrated coefficients to project (predict) a readiness index for each student, measured on a 0-100 scale. Readiness is also reported as the projected probabilities of high school graduation status. In this report, the readiness index uses predictors from 8th grade and the first year in high school. However, the system could readily be extended to include additional student outcomes from subsequent years and grades, yielding readiness projections with increasing accuracy. Our system reports both overall readiness indices, components of readiness, and high school graduation probabilities to enable diagnosis of student and school-level strengths and weaknesses at the beginning of the school year.

Second, we develop equity-focused measures of student and school-level readiness that highlight differences in readiness both within and across schools by multiple student characteristics, including poverty status, race/ethnicity, and other demographic and program participation indicators. We report metrics for each demographic group of interest within the school, which allows us to identify differences and gaps between them. In addition, we report metrics that enable comparisons of the performance of different demographic groups within the school with the performance of demographic groups in schools with similar characteristics.

Third, we develop retrospective measures that show how students' readiness changed from one year to the next. This will become that basis for measuring school performance and holding schools accountable for improving readiness. The goal is to measure schools' contribution to student readiness. School performance is assessed using the multiple measures of readiness, in contrast to standard accountability models that focus only on math and English test scores. These metrics are reported at the end of the school year. Schools' performance in improving readiness is summarized by a school progress index, which is created as follows:

- We regress the readiness index described above on 8th grade predictors (the same four as before: attendance, course enrollment, test scores, and course GPAs) and an indicator for which high school the student attended. The estimated school effect measures average student progress with respect to overall student readiness. The model used is a standard multilevel model with fixed school effects (Wooldridge, 2010; Raudenbush and Bryk, 2001; Snijders, 2011).[1]
- The above method is also applied to each of the four readiness components to produce separate school effects for each component.
- Finally, statistical shrinkage is applied to school progress metrics to adjust for estimation error.

Fourth and finally, we construct equity-focused measures of school quality. Whereas standard accountability models measure school quality as a unidimensional construct, it is essential to extend these measures to allow for the possibility that school quality is not the same for all students but rather differs by student characteristics. This part of the system parallels the equity-aligned metrics of student status but focuses on how school quality differs for students with different student characteristics and across schools with different student compositions.

Defining Consequential Outcomes & Input Variables

In this report, we focus on building analytics around one highly valued medium-term outcome: high school graduation.[2] We allow for several different high school outcomes:

- dropout or non-graduate
- regular (non-honors) diploma
- honors (or advanced) diploma

---

[1] This model is also used to construct readiness indices using 8th grade outcomes, included as predictors in the model of 9th grade readiness.

[2] We have explored incorporating long-term outcomes into EAAS, including college enrollment and completion and career and workforce success, but limit our focus in this report to focusing on high school graduation as the key student outcome. We believe that expanding the system to incorporate multiple long-term outcomes and multiple pathways to long-term success could be important from an equity perspective and discuss this at the end of the report.

Short-term K-12 outcomes that go beyond test scores are another critical piece of the system outlined here. In this report we demonstrate implementation of the system using four types of student data, but states and school districts that adopt this system could choose to use different predictors that they find valuable from a policy standpoint. **Table 1** below summarizes the data included in the current system.

Table 1. Short-Term Outcome Types & Inclusion Reasons

Variable Included	Reason for Inclusion
Annual student attendance data	Attendance can be used as a proxy for a student's participation in their education and measures whether a student had the opportunity to learn in their classroom.
Enrollment in challenging courses (all subjects)	Students' opportunity to participate in challenging coursework demonstrates mastery to colleges and poses a serious challenge to equity as students continue to be "tracked" based on demographics.
Test scores in end-of-course (EOC) and standardized statewide assessments[3]	End-of-course exam outcomes are potentially strong predictors of long-term outcomes because they measure multiple outcomes: • Did the student take a course (and in what grade)? • Did the student take the EOC assessment (and how many times)? • How did the student perform on the test, particularly if passing the test is required for graduation, as required in some states?
Grade point average (GPA)	Course grades have been and continue to be used in college admission decisions and have been the key (and sometimes only) outcome measure used in early warning systems.

[3] Certification exam scores are also available for numerous CTE fields and could be included as either short-term or long-term workforce outcomes. Additionally, Advanced Placement (AP) and International Baccalaureate (IB) could be included, particularly in models that include outcomes in higher grades.

Figure 2. Diving Deeper into the Structure of the System

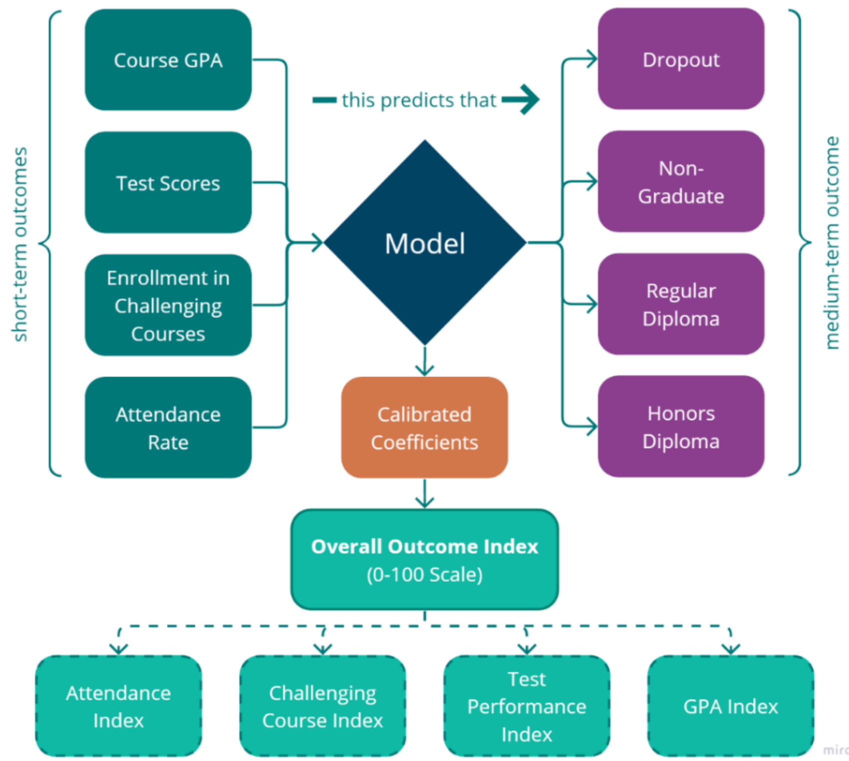


Figure 2 combines the steps explained in the previous section with the more granular understanding of the consequential outcomes provided in this section. Four categories of predictors ("short-term outcomes") are used to predict a categorical high school graduation outcome ("medium-term"). The coefficients may then be used to generate the composite "Overall Readiness Index" or decomposed into readiness indices for each of the predictor categories.

A key objective of school accountability systems is to incentivize schools and districts to improve student and school performance on the dimensions included in the system. There is a strong consensus that it is more appropriate to incentivize behavior when evaluation is based on multiple indicators to avoid narrow teaching to the test *(**MET Policy & Practice Brief**)*. By shifting to a broad set of student outcomes, it may be more difficult to "game" this system and teach narrowly to the set of outcomes.

One concern about including course grades in an integrated early warning and school accountability system is that this could spur grade inflation.[4] We address this problem by

[4] It is also possible that including course enrollment variables as a student outcome could encourage schools to enroll students in advanced courses even if they are not prepared to succeed in these courses. From an equity standpoint, we are interested in providing incentives to schools to offer challenging course and to enroll disadvantaged students in these courses.

adjusting grade variables to eliminate school differences that are not predictable given prior student information but maintain within-school differences. Put differently, our selected method for grade adjustment allows for variation across schools if that is justified by student differences in the 8th grade predictors. An important advantage of this approach is that it allows us to retain grades in the integrated system, typically one of the strongest predictors of readiness, without creating adverse incentives. On the other hand, our models of school quality remove the positive effects of having students with high performance to start with, as this would unfairly advantage the high school compared with a high school that received students with low prior performance in 9th grade. The school quality model only retains the portion of the positive trend that cannot be explained by 8th grade performance. **Appendix B** provides technical information and statistical results for the grade adjustment method.

III. HOW EAAS WOULD LOOK IN PRACTICE

In this section, we describe the key parts of the Equity-Aligned Analytics System and illustrate how the educator, parent, and student-facing reports we envision would look using real-world student data. Technical details on the models and statistical methods used to implement the EAAS and estimates of these models are presented in Technical Appendices. All the analytics and statistical results presented in this report are based on real-world data. Summary statistics for this data are provided in the relevant sections and appendices.

A Large, Diverse American District[5]

We worked with a large metropolitan district to hone their accountability system and generate reports that would illustrate the additional components that make it more robust, and equity aligned. We obtained a panel of student data for demographics, attendance, course-taking, test scores, and grades that extended from elementary through high school. We also included graduation outcomes after four years and used all these data points to generate sample input sets to test our models' efficacy. Once we had model outputs, we worked to design preliminary reports that can clearly communicate student readiness and school quality to various audiences.

In the next two sections we present the proposed application of the central parts of the Equity-Aligned Analytics System (EAAS) in the real world. The first section shows how the measures of projected readiness are constructed and how they are presented in student, school, and district reports. In the second section, we show how measures of student progress on high school outcomes and projected readiness are constructed and presented in reports that can be used in an expanded and equity-aligned school accountability system.

---

[5] In the report we do not identify the state and district that generously provided this project with extensive data on students in 8th grade through college graduation, per the administration's wishes.

Prospective Analytics: Early Warning Analytics

In this section we first describe the steps required to construct projections of future readiness, defined in terms of high school graduation status. Then we present and discuss the reports constructed by using the readiness metrics.

We have selected four types of student outcomes to use in measuring students' readiness with respect to high school graduation status: attendance rate, enrollment in advanced math and science courses, scores on high school end-of-course exams (measured on a 0-to-100 point scale where 65 is the threshold for passing), and course grades in the following subjects: math, science, English, social studies, and all other subjects (measured on the standard A-to-F scale).[6]

This set of outcomes can be viewed as an expanded student report card in that it includes outcomes other than grades. **Table 2** provides examples of the outcomes for four different 9th grade students, listed as: (1) High Readiness, (2) Medium Readiness, (3) Low Readiness, and (4) Very Low Readiness. The high-readiness student had an attendance rate of 99%, was enrolled in advanced courses in both math and science, took two end-of-course exams (geometry and chemistry) and earned high scores on both exams, received A grades in all courses other than English, and received a grade of B in English. In contrast, the very low-readiness student had an attendance rate of 78% (an example of chronic absenteeism), was not enrolled in advanced courses in math or science, took one end-of-course exam and received a non-passing score of 45, and received a mixture of C and D grades. The medium and low-readiness students had outcomes worse than the high-readiness student, but better than the very low-readiness student.

[6] In the district used for this study, the offerings of advanced (honors) versus regular (non-honors) courses were primarily available in math and science for high school year 1 (9th grade) students. Algebra 1 was considered a regular math course and geometry was considered an advanced course at that grade level.

Table 2. Student Report Card Information for Example Students in 9th Grade

Outcome	Example Student 1	2	3	4
	High Readiness	Medium Readiness	Low Readiness	Very Low Readiness
Attendance	99	92	88	78
Advanced Courses				
Math	Yes	Yes	No	No
Science	Yes	Yes	No	No
Test Scores				
Geometry	85			
Chemistry	92			
Biology		67		
Algebra 1		78	45	40
Course Grades				
Math	A	A	D	F
Science	A	B	C	D
English	B	C	C	C
Social Studies	A	B	B	D
Other subjects	A	C	C	C

We translate the raw student report card information, as displayed in **Table 2**, into actionable readiness data through the following steps:

Step 1: Identify the medium-term outcomes to predict. We chose high school graduation as the equity-oriented outcome to predict because graduation is typically the gateway to future opportunities that determine lifelong well-being. This district's system allowed for three levels of the outcome: 1) Did not graduate within four years; 2) regular diploma; 3) Honors diploma. One of the primary advantages of using an outcome variable with this structure is that it embeds in a very simple way two levels of high school performance since the requirements for an honors diploma are higher than for a non-honors diploma. Since students with honors diplomas tend to obtain higher levels of postsecondary education than students with non-honors diplomas, the high school graduation outcome used as our future outcome measure incorporates both the medium-term outcome of high school graduation and, implicitly, the long-term outcome of college enrollment and graduation.[7] The model details and the associated formulas are presented in **Appendix A**.

[7] In some states, an honors-type diploma is required for admission to public four-year colleges. For example, in California, the so-called AG diploma is a requirement for admission to a California State University or University of California institution (Freshman Application Guide (calstate.edu)). In other cases, receipt of an honors diploma may increase the probability of admission to a selective college.

Step 2: Choose the statistical model. Given that the medium-term outcome variable to be predicted is a discrete, multi-valued, and ordered variable, the best statistical model to realistically represent this outcome is an ordered probit or logit model (Maddala, 1983; Daykin & Moffatt, 2002; Greene, 2017; Woodridge, 2010). Although both models yield nearly identical results, we use the ordered probit model because it is most compatible with the regression models used elsewhere in this report.

This approach has several advantages. First, it is straightforward to include multiple student outcomes in the model. Second, the model produces calibrated coefficients that are used to weight the multiple student outcomes to produce best predictions of overall (total) readiness and readiness for separate components (sets of variables). Calibrated coefficients can be estimated (updated) annually (or with less frequency, as needed) to best represent the possibly changing predictive relationships between student outcomes and medium (and long-term) outcomes. Third, it is straightforward to build the model to allow the relationships between predictors and high school graduation status to be non-linear. Fourth, the model is explicitly designed to take account of the fact that high school graduation status is a discrete, ordered outcome and, as a result, it is straightforward to calculate the projected probabilities of each high school graduation outcome, given projected readiness.

Step 3: Choose high school outcome/predictors and model design. Given the fact that the number of high school outcome variables included in EAAS is large, motivated by the desire to substantially expand the number of outcomes included in both prospective and retrospective, we prioritized the following criterion in defining the predictors: (1) Define the predictors to be as simple as possible without substantially sacrificing the predictive power of the model; (2) Allow for the possibility that the effects of attendance and test scores may vary at different levels of these variables, including at policy-relevant values: whether a student was chronically absent from school and whether a student passed or failed an end-of-course exam.

After applying these criteria, we defined the variables for each of the four outcome components as follows:

Description	Indicator Variables
Student attendance rate A set of categorical indicator variables that represent the attendance levels.	*-Very high chronic absenteeism (baseline value) (<81%)* *- Chronically absent (81-90%)* *- Low attendance (91-93%)* *- Medium attendance (94-96%)* *- High attendance (97-98%)* *- Very high attendance (99-100%)*
Advanced course enrollment Separate indicator variables for each subject.	*- Not enrolled (baseline value)* *- Advanced math (geometry or Algebra 2)* *- Advanced science*
End-of-Course Test Score Variables Counts of test scores in each score band, summed across courses.	*Test score (baseline value) (0-49)* *Test score (50-59)* *Test score (60-64)* *Test score (65-75)* *Test score (76-85)* *Test score (86-100)*
Course/subject grade point average (GPA) by subject[8] Separate GPA variables graded on the standard 4.0 scale (A = 4, B = 3, C = 2, D = 1, F = 0).	*Math* *Science* *English* *Social studies* *Other courses/subjects*

Step 4: Model calibration. We use historic data on both high school outcomes and the future student outcome to calibrate (or estimate) the model parameters to be used in constructing new readiness projections of future outcomes. We employed the data of the cohort that graduated from high school in the 2017-18 school year. Since four years of data are required to calibrate the high school graduation model, calibration coefficients based on pre-COVID data are likely to be the preferred coefficients until at least the end of the 2024-25 school year if it is desirable to avoid using COVID period data to calibrate the high school graduation models.

The estimated calibration coefficients are reported in **Table A1** in **Appendix A**. These coefficients have been transformed so that projected readiness values range from 0 to 100. As indicated in the table, the 9th grade outcomes are all very strong predictors of high school graduation status and, as a result, the explanatory power of the model is strong: the Pseudo R-squared statistic equals 0.73. This statistic is high, which indicates that student performance in 9th grade is a very strong early warning signal, but not so high as to preclude the possibility

[8] As discussed earlier in the report, the GPA variables for each subject area are adjusted to eliminate any incentive for schools to artificially inflate grades. A more detailed explanation of the method is presented in **Appendix B**.

that it is possible to improve students' likelihoods of graduating from high school with either a non-honors or honors diploma.

As expected, the effects of high student attendance and achievement on the end-of-course tests are strongly increasing with effects equal to approximately 12 points for the highest level of attendance and a test score of 86 or greater for a single exam outcome. The baseline effects for these two outcomes, an attendance rate less than 81% and a test score less than 50 (or no test score), are anchored at zero. Since many students take two (and sometimes more) end-of-course tests each year, the potential effects of test score performance could be two to three times the effect of a single exam score. The effects of enrollment in an advanced math or science course are strong and approximately equal (5 points) and the combined effect of enrollment is double that amount to 10 points. Grade (GPA) effects in all subjects are strong, but especially strong in math and English.

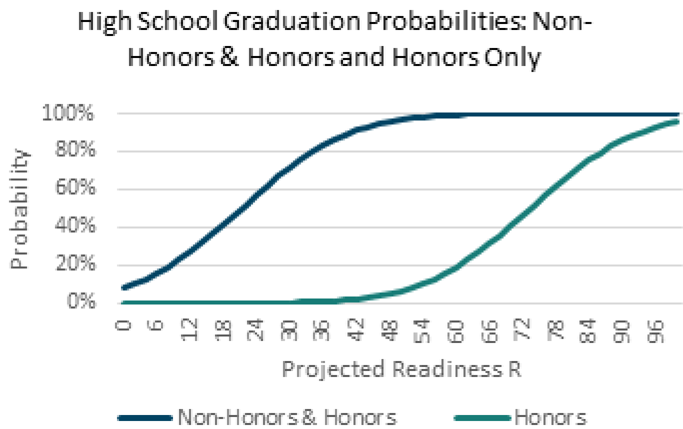
Step 5: Construction of projected readiness for all students. Given calibrated coefficients and data on 9th grade outcomes, it is straightforward to construct projected readiness metrics, overall and by the four components: attendance, advanced course enrollments, end-of-course test scores, and grades by subject. Projected readiness values for each student are obtained by multiplying the calibrated coefficient with the corresponding student outcomes and then summing the products. The projected readiness values range from 0 to 100 points.

We label the readiness index with letter R and the components of the index for attendance, advanced courses, tests, and grades as R1, R2, R3 and R4, respectively. Breaking the overall index into separate components, allows students, parents, and educators to target the areas where students have weaknesses, while building on areas of strength.

One of the primary advantages of reporting projected readiness as described above is that the overall (total) index is approximately continuous and, as demonstrated later in the report, strongly differentiates readiness levels for different students along the entire spectrum of readiness. However, despite the advantages of reporting projected readiness on a continuous scale, the perceived validity of this metric is likely enhanced by showing how the metric is related to projected probabilities of high school graduation outcomes: non-graduation, graduation with a non-honors diploma, and graduation with an honors diploma. These latter outcomes are more tangible and less abstract than the projected readiness indices and other outcomes commonly included on school report cards and in accountability systems (such as test scores).

To obtain some insight into the similarities and differences between the continuous readiness metric R and the readiness probabilities, **Figure 3** graphs the probabilities of graduating with: **(1) a non-honors or honors diploma (dark-blue line)** and **(2) an honors diploma (dark-green line)**. These probabilities are functions of readiness metric R. The probability of graduating with the higher diploma is always lower than the probability of graduating with either diploma, except at the extremes of the range of R values, where the probabilities equal zero or one, respectively. Both lines are highly nonlinear and do not differentiate students at both ends of the range of projected readiness R. The bottom line is that the high school graduation probabilities are connected to tangible events and therefore have face validity, but they are not as useful as the projected readiness index R in differentiating student readiness. As a result, the example early warning reports use both readiness metrics.

Figure 3. Non-Honors and Honors High School Graduation Probabilities Given Projected Readiness *R*



Early Warning Reports

After applying the steps presented above, the raw student report cards shown at the beginning of the section get enriched with new actionable readiness metrics for each student, as shown in **Table 3**. The reports include the same information as contained in the expanded school report cards in, but they also include projected readiness values for each component and the associated high school graduation probabilities. The latter metrics are available because the calibrated high school graduation status model, trained on historical data, is used to project (predict) these future outcomes given student outcome data known in 9^{th} grade. These reports are intended for use as early warning reports for students entering 10^{th} grade. The reports provide information on:

1. Student outcomes in all four areas: attendance, advanced course taking, test scores, and course grades (the same data reported in Table 2)[9]
2. Readiness points for high school graduation status on the 0-to-100 scale for each of the four components (R1 to R4) and the total outcome (R)
3. A readiness label (with color coding), determined by the total projected readiness outcome (*R*):
   a. **Q1: Very Low Readiness**: Readiness Index R < 25 (code red)
   b. **Q2: Low Readiness**: 25 < Readiness Index R < 50 (code orange)

[9] The reports provide space for up to two end-of-course test scores. The course exam taken by the student is listed. These vary across students. If a student did not take an exam, an “na” is listed.

c. **Q3: Medium Readiness**: 50 < Readiness Index R < 75 (code green)
d. **Q4: Very High Readiness**: 75 < Readiness Index R (code blue)[8] As discussed earlier in the report, the GPA variables for each subject area are adjusted to eliminate any incentive for schools to artificially inflate grades. A more detailed explanation of the method is presented in Appendix B.

4. Projected high school graduation probabilities for the outcomes: non-honors diploma or higher and honors diploma.

The thresholds used to assign students to readiness status levels are values that should be set by the schools, districts, or states since they are intended to be trigger action to support students with lower readiness. For simplicity, we selected threshold values of 25, 50, and 75 on the 0/100 scale. In the sample used to calibrate the high school graduation model, this results in the following distribution of students: 15.4% with R below 25, 30.5% with R between 25 and 50, 31.7% with R between 50 and 75, and 22.5% with R above 75. Raising the threshold for the very low readiness level would necessarily increase the proportion of students assigned to that category if it was desirable to identify a larger proportion of students at that level.

The four early warning student reports in **Table 3** represent students at very different levels of the distribution of projected readiness, with total readiness scores equal to 91, 68, 29, and 15 on the 0/100 scale. It is interesting to contrast the reports for the students with low and very low readiness, with total projected readiness points of 29 and 15 points, respectively, a difference of 14 points. This difference may seem small, but the probability of graduating from high school with a non-honors diploma for the low readiness student is twice as large (69%) as the probability for the very low readiness student (34%).

Table 3. Student Report Card Information for Example 9th Grade Students

	Example Student							
	1		2		3		4	
	High Readiness		Medium Readiness		Low Readiness		Very Low Readiness	
	Outcome	**Readiness**	**Outcome**	**Readiness**	**Outcome**	**Readiness**	**Outcome**	**Readiness**
Constant		1		1		1		1
Attendance	99	**12**	92	**7**	88	**5**	78	**0**
Advanced Courses		**10**		**10**		**0**		**0**
Math	Yes		Yes		No		No	
Science	Yes		Yes		No		No	
Test Scores		**24**		**16**		**0**		**0**
Geometry	85							
Chemistry	92							
Biology			67					
Algebra 1			78		45		40	
Course Grades		**44**		**34**		**23**		**14**
Math	A		A		D		F	
Science	A		B		C		D	
English	B		C		C		C	
Social Studies	A		B		B		D	
Other Subjects	A		C		C		C	
Total Readiness		**91**		**68**		**29**		**15**
Readiness Status	R>75	**Q4: Very High**	50<R<75	**Q3: Medium**	25<R<50	**Q2: Low**	R<25	**Q1: Very Low**
Post High School Probabilities:								
Graduate: non-honors diploma		100%		100%		69%		34%
Graduate: honors diploma		88%		35%		0%		0%

What specifically accounted for this large difference? Both students earned zero points by not taking an advanced course and zero points because they failed to pass an end-of-course exam. Both students were chronically absent, but the absenteeism rate was much lower for the very low performance student (78% versus 88%), resulting in 5 fewer points. The other major difference was in course grades. The low readiness student received grades of B, C and D, earning 23 points, whereas the very low readiness student received grades of C, D, and one F, earning 15 points. This early warning data should trigger support (or interventions) for both students, especially for the very low readiness student, with a focus on student attendance and instructional support in the students' weakest subjects: math for the low readiness student and math (especially) and science and social studies for the very low readiness student.

The following early warning reports in this section consider how student early warning data could be summarized in classroom, school, district, or state reports. The prime objective in early warning reports is to present information on projected student readiness to spur teachers, principals, district, and state educators to intervene and provide support if students are not on track to graduate from high school. For simplicity, we present two types of reports. The first type includes student data for all students in a classroom or school and aggregates that data to the next level, a classroom or school level. We refer to this type of report as a school report and it is designed to provide individual and aggregate data for educators who are charged with directly supporting students. The second type includes classroom, school, or district-level data and aggregates that data to the next level, a school, district, or state. We refer to this type of report as a district report and it is designed to provide school-level and district-wide data relevant for managing resources and setting policies to support teaching and learning.

Before discussing the school and district reports in more detail let's first introduce another readiness metric which is used in the reports to assign the school readiness level. An important feature of the school-level tables is identifying the readiness status of schools (the parallel measure to classifying student readiness levels). The thresholds used to assign schools to readiness status levels are values that should be set by the schools, districts, or states since they are intended to trigger action to support students and schools' lower readiness. In this report we have defined school readiness levels based on the proportion of students at student readiness level Q1 (R <= 25). After inspecting the district-wide cumulative distribution of students at readiness level Q1, we set the following threshold values for five school readiness levels, designated S1 to S5:

Proportion of Students in Readiness Categories					
School Readiness level	**S1: Very Low Readiness**	**S2: Low Readiness**	**S3: Medium Readiness**	**S4: High Readiness**	**S5: Very High Readiness**
Thresholds on Q1 Proportion for Classifying Readiness	**27 - 55**	**18 - 26**	**10 - 17**	**3 - 9**	**0 - 2**
Proportion of Students at Readiness Level	17.8%	22.4%	19.4%	24.5%	15.8%

Given the selected threshold values, 17.8% of schools (weighted by the number of students) are classified with very low readiness and 22.4% are classified with low readiness. Raising the threshold values for the very low readiness level would necessarily increase the proportion of students assigned to that category if it was desirable to identify a larger proportion of students at that level. The school readiness level designation is included in all school tables.

Early warning school reports are presented in **Tables 4(A, B)**, **5(A, B)** and **6(A, B)** for the same student cohort of students represented in **Table 3**. **Tables 4A**, **5A**, and **6A** present the three parts of the early warning report which show the results for a school (School A) with a high concentration of students with high readiness. **Tables 4B**, **5B**, and **6B** present the three parts of the early warning report which show the results for a school (School B) with a high concentration of students with low readiness. These reports present key readiness information on both individual students and schools (in the aggregate) and could be the natural point of entry for teachers and school leaders since they characterize the degree to which schools in each grade have large numbers of students with low readiness to graduate and explicitly identify the readiness status of all students.

The top panel (**Panel 1**) of the early warning school report shown in **Tables 4A** and **4B** contains information on overall readiness for each student in the school: the total readiness index R, the two high school graduation status probabilities, and the readiness status. The panel includes information on an abbreviated list of students in the school to limit the size of the table in this report. Additional information for each student could be provided in expanded tables, particularly if the data is available via flexible online software.

Panels 2 and **3** summarize student readiness data at the school level. **Panel 2** presents average school-level outcomes: projected readiness R and the high school graduation proportions. **Panel 3** focuses not on average readiness but rather on the distribution of readiness across all students; the proportions of students at each of the four readiness levels, Q1 to Q4, are reported.

Table 4A. School Early Warning Report, High Readiness School (School A)

School Early Warning Report – Part 1				
School A: High Readiness School				
Grade: 9				
Panel 1: Metrics by Student				
	Readiness (R)		**Graduation Probabilities**	
Student ID	**Readiness Status**	**Total Index**	**Non-honors +**	**Honors**
1	Q4: Very High	91	100%	88%
2	Q3: Medium	68	100%	35%
3	Q2: Low	29	69%	0%
4	Q1: Very Low	15	34%	0%
…	*Student list abbreviated in this table.*			
Panel 2: Average School Statistics				
	Readiness (R)		**Graduation Proportions**	
School Size	**School Average Index**		**Non-honors +**	**Honors**
1151	83		99%	70%
Panel 3: Proportion of Students at Readiness Levels				
School Readiness Level	**Q1: Very Low Readiness**	**Q2: Low Readiness**	**Q3: Medium Readiness**	**Q4: Very High Readiness**
S5: Very High	1%	2%	19%	79%

Table 4B. School Early Warning Report, Low Readiness School (School B)

School Early Warning Report – Part 1				
School B: Low Readiness School				
Grade: 9				
Panel 1: Metrics by Student				
	Readiness (R)		**Graduation Probabilities**	
Student ID	**Readiness Status**	**Total Index**	**Non-honors +**	**Honors**
1	Q3: Medium	62	100%	23%
2	Q2: Low	41	90%	2%
3	Q2: Low	27	64%	0%
4	Q1: Very Low	8	19%	0%
...	*Student list abbreviated in this table.*			
Panel 2: Average School Statistics				
	Readiness (R)		**Graduation Proportions**	
School Size	**School Average Index**		**Non-honors +**	**Honors**
348	45		71%	15%
Panel 3: Proportion of Students at Readiness Levels				
School Readiness Level	**Q1: Very Low Readiness**	**Q2: Low Readiness**	**Q3: Medium Readiness**	**Q4: Very High Readiness**
S2: Low	21%	38%	32%	9%

The differences in readiness statistics between schools A and B are enormous. **Panel 2** shows that the school average readiness index is much higher in school A than in school B, 83 and 45 respectively. **Panel 3** shows that the proportion of students with very low readiness projections (Q1) is 1% in School A and 21% in School B and the proportion of students with very high readiness (Q4) is 79% in School A and 9% in School B. School A is thus designated as a school with very high readiness (S5) and school B is designated as a school with low readiness (S2). In school A the projected proportion of students graduating with either diploma is 99% and the projected proportion of students graduating with an honors diploma is 70%. In contrast, in school B the projected proportion of students graduating with either diploma is 71% and the projected proportion of students graduating with an honors diploma is 15%.

Tables 5A and **5B** present **Panel 4** of the school early warning report which summarizes student readiness data at the school level for two demographic groupings: economically disadvantaged (ED) (yes/no) and race/ethnicity (black, Hispanic, and other). These tables could be expanded to consider data for additional demographic groups. In School A the proportion of students at the very low readiness level is only 1% or lower. Therefore, the readiness status for all demographic groups in this school is very high (school level S5) but with slightly lower average projected readiness for black and Hispanic students. In School B school readiness level is low (S2) when all students are aggregated, but there are differences across demographic groups with economically disadvantaged, Hispanic, and other race/ethnicity students having a lower readiness level compared to the other comparable subgroups.

Table 5A. School Early Warning Report, High Readiness School (School A)

School Early Warning Report – Part 2						
Panel 4: Readiness by Demographic Groups						
Group	**Number of Students**	**Group Proportion**	**Readiness Status**	**Average Readiness**	**Proportion of Students**	
					Q1: Very Low Readiness	**Q2: Low Readiness**
All Std.	1151	100.0%	S5: Very High	83.0	1%	2%
Economically Disadvantaged:						
No	771	67.0%	S5: Very High	83.0	0%	2%
Yes	380	33.0%	S5: Very High	82.9	1%	2%
Race/Ethnicity:						
Black	76	6.6%	S5: Very High	78.0	1%	4%
Hispanic	73	6.3%	S5: Very High	77.6	1%	7%
Other	1002	87.1%	S5: Very High	83.7	0%	1%

Table 5B. School Early Warning Report, Low Readiness School (School B)

School Early Warning Report – Part 2						
Panel 4: Readiness by Demographic Groups						
Group	**Number of Students**	**Group Proportion**	**Readiness Status**	**Average Readiness**	**Proportion of Students**	
					Q1: Very Low Readiness	**Q2: Low Readiness**
All Std.	348	100.0%	S2: Low	44.8	21%	38%
Economically Disadvantaged:						
No	172	49.4%	S3: Medium	47.8	15%	40%
Yes	176	50.6%	S1: Very Low	41.8	28%	36%
Race/Ethnicity:						
Black	165	47.4%	S3: Medium	48.3	13%	41%
Hispanic	160	46.0%	S1: Very Low	41.5	28%	36%
Other	23	6.6%	S1: Very Low	42.2	30%	30%

Panels 5 and **6** in **Tables 6A** and **6B** return to a focus on student data within each school and provide data on projected readiness for all four components: attendance, advanced course enrollment, test scores, and grades/GPA. These tables are no substitute for the more complete student reports (shown in **Table 3**) but provide a compact schoolwide perspective on student strengths and weaknesses with respect to the four readiness components. This could help educators craft actions that are responsive to individual student needs. We signal the need for this step by including spaces in the table for teachers to log planned actions for each student and for the school. Expanded versions of these early warning tables could provide lists of resources, programs, and interventions that could be provided to students based on their individual needs.[10]

[10] Expanding an early warning system to collect information on actions spurred by the system would be highly valuable since it could potentially be used to evaluate the impact of these actions.

Table 6A. School Early Warning Report, High Readiness School (School A)

School Early Warning Report – Part 3							
Panel 5: Projected Student Readiness (Total and by Component) and Planned Actions							
Student ID	**Readiness Status**	**Total Readi-ness**	**Attend-ance**	**Advan-ced Courses**	**Test Scores**	**Grades/ GPA**	**Planned Actions for Each Student**
1	Q4: Very High	91	13	10	24	44	*List actions*
2	Q3: Medium	68	8	10	16	34	
3	Q2: Low	29	6	0	0	23	
4	Q1: Very Low	15	1	0	0	14	
…	*Student list abbreviated in this table.*						

Panel 6: Planned Schoolwide Actions on Overall Readiness and by Component				
Total Readiness	**Attendance**	**Advanced Courses**	**Test Scores**	**Grades/ GPA**
List actions				

Table 6B. School Early Warning Report, Low Readiness School (School B)

School Early Warning Report – Part 3							
Panel 5: Projected Student Readiness (Total and by Component) and Planned Actions							
Student ID	**Readiness Status**	**Total Readi-ness**	**Attend-ance**	**Advan-ced Courses**	**Test Scores**	**Grades/ GPA**	**Planned Actions for Each Student**
1	Q3: Medium	61	10	5	16	30	*List actions*
2	Q2: Low	40	8	0	6	26	
3	Q2: Low	27	6	0	0	20	
4	Q1: Very Low	10	1	0	0	9	
…	*Student list abbreviated in this table.*						

Panel 6: Planned Schoolwide Actions on Overall Readiness and by Component				
Total Readiness	**Attendance**	**Advanced Courses**	**Test Scores**	**Grades/ GPA**
List actions				

District Reports

Although district and state educators should have access to school reports in well-designed early warning systems, it is useful to design reports and early warning data systems and dashboards that make it easy for educators to access and digest this information. **Tables 7** and **8** present examples of district reports.

Table 7. District Early Warning Report, Part 1

District Early Warning Report / School Data						
District D1						
Grade: 9						
Panel A1					**Grad. Diploma Proportions**	
School ID	**School Size**	**School Readiness Level**	**Projected School Readiness**		**Non-honors +**	**Honors**
A	1151	S5: Very High	83		99%	70%
B	348	S2: Low	45		71%	15%
...	School list abbreviated in this table.					
Panel A2			**Proportion of Students at Readiness Levels**			
School ID	**School Size**	**School Readiness Level**	**Q1: Very Low Readiness: R Index < 25**	**Q2: Low Readiness: R Index = (25,50)**	**Q3: Medium Readiness: R Index = (50,75)**	**Q4: Very High Readiness: R Index > 75**
A	1151	S5: Very High	1%	2%	19%	79%
B	348	S2: Low	21%	38%	32%	9%
...	*School list abbreviated in this table.*					
Panel A3	**Proportion of Students at Very Low Readiness Level (Q1)**					
School ID	**School Size**	**Economically Disadvantaged**		**Race/Ethnicity**		
		No	**Yes**	**Black**	**Hispanic**	**Other**
A	1151	S5: 0%	S5: 1%	S5: 1%	S5: 1%	S5: 0%
B	348	S3: 15%	S1: 28%	S3: 13%	S1: 28%	S1: 30%
...	School list abbreviated in this table.					

The reports follow roughly the same design as the school reports in that **Table 7** reports average school readiness data for all schools, compiled from Tables 4 and 5 above. Table 8 presents districtwide data. In **Table 7**: (a) **Panel A1** reports average projected readiness and the proportion of students by diploma status, (b) **Panel A2** reports the proportion of students in each student readiness category (Q1 to Q4), and (c) **Panel A3** reports the proportion of students at the very low readiness level (Q1) by demographic group. **Table 8** reports the same data in **Panels B1** to **B3** aggregated to the district level.

Table 8. District Early Warning Report, Part 2

District Early Warning Report / Average District Data					
District: D1					
Grade: 9					
Panel B1				Grad. Diploma Proportions	
		Projected Readiness R		Non-honors +	Honors
		52.8		85%	26%
Panel B2	Proportion of Schools at Readiness Levels				
Number of Schools	S1: Very Low Readiness: Proportion >= 27	S2: Low Readiness: Proportion = (18,26)	S3: Medium Readiness: Proportion = (10,17)	S4: High Readiness: Proportion = (3,9)	S5: Very High Readiness: Proportion <= 2
>25	17.8%	22.4%	19.4%	24.5%	15.8%
Panel B3					
Group	Readiness Status	Average Projected Readiness R	Q1: Very Low Readiness: R Index <= 25	Q2: Low Readiness: R Index = (26,50)	Group Proportion
All Students	S3: Medium	52.8	15.4%	30.5%	100.0%
Econ. Disadv.					
No	S3: Medium	58.9	9.7%	26.1%	52.1%
Yes	S2: Low	46.2	21.5%	35.2%	47.9%
Race/Ethnicity					
Black	S2: Low	43.4	23.1%	39.2%	26.0%
Hispanic	S2: Low	45.3	19.8%	38.4%	38.4%
Other	S4: High	67.9	5.0%	15.6%	35.6%

Effective school, district, and state leaders might engage with early warning reports of this type and, preferably, flexible early warning data systems to address the questions listed below. Engagement with this data is intended to spur actions by teachers and leaders to support schools with high concentrations of low or very low readiness students and to thereby remediate systemic differences in projected readiness across different demographic groups.

1. **What is the overall readiness challenge?** The districtwide proportion of students with very low readiness status is 17.8%. **How many students are expected to graduate from high school?** 85%. **To graduate with honors degrees?** 26%.
2. Equity challenge: **Do readiness levels vary systemically across demographic groups?** Yes, the proportion of low readiness students is much higher (level S2) for economically disadvantaged students and black and Hispanic students (20% – 23%).
3. **Which components are the source of low readiness: attendance, advanced course taking, test scores, or grades?**
4. Dig deeper: Sort the list of schools by readiness and identify all schools at the very low readiness level (S1).

5. Equity inquiry: For demographic groups flagged in inquiry #2, sort the list of schools by readiness status for each low-readiness group and identify all schools at the very low readiness level (S1).
6. School-by-school deeper dive: For schools identified as very low readiness schools in inquiries #4 and #5, drill down to the detailed school reports. **Which schools are especially challenged with respect to each component?**

Retrospective Analytics: School Progress and Accountability

In this section we describe how the early warning measures can be used in an expanded and equity-aligned school accountability system and, more generally, can be used by educators to reflect upon progress made and the drivers of change. The basic idea is to use the change in readiness for individual students and the average progress made by students at the school and district levels from one school year to the next. Since the retrospective progress metrics are based on the prospective early warning metrics, they show stakeholders whether the actions taken based on the early warning metrics succeeded in increasing student readiness.

The progress model is a retrospective model, and it is like a traditional growth or value-added model in that it aims to estimate the contribution of schools to outcomes measured at the end of high school year 1 (9^{th} grade) – the post-year outcomes – controlling for student outcomes measured at the end of 8^{th} grade, the prior (pre-year) outcomes. Whereas growth models typically focus on single student outcomes such as math and English language arts (ELA) achievement, the dependent variables in this model – projected readiness or readiness components – are composite variables comprised of the multiple high school outcomes that are included as predictors in the high school graduation status model. Similarly, whereas growth models typically include a limited set of student outcomes as predictors or control variables (for example, prior test scores in math and/or ELA), the EAAS progress model includes a set of 8^{th} grade variables that fully match the four components included in the high school graduation status model. The technical details of the progress models are described in **Appendix C**. The resulting progress metrics would be added to school and district accountability systems. The methodology used to develop the progress metrics involves the following steps:

Step 1: Estimate student readiness for grade 8. In the previous section we presented readiness metrics calculated for each student by using 9^{th} grade predictors. We will refer these metrics as 9^{th} grade readiness metrics. We use the 9^{th} grade readiness metrics as dependent variables in a linear model with 8th grade controls and school fixed effects to calibrate the coefficients for 8^{th} grade variables. Then we employ the coefficients to predict 9^{th} grade readiness from 8th grade variables. This procedure is applied for overall readiness as well as for each component.

Another approach is to follow a similar procedure to calculate the 8th grade readiness as we did for 9th grade readiness demonstrated in the previous section. We implemented this alternative approach to assess the robustness of our main approach and it yielded very similar results.

Note the coefficient for 8th grade variables are estimated every year as part of the process to generate the progress metrics. As such these coefficients can be used to predict readiness for the next cohort entering the high school and will be reported as part of the early warning reports delivered at the beginning of the year. **Table C2** in **Appendix C** reports the calibrated coefficients which are used to generate the 8th grade readiness metrics we present here.

In the report tables presented later in this section, we denote the prediction of 9th grade readiness by 8th grade predictors with the letter P. The 9th grade readiness metrics itself are denoted with the letter R. Metrics P and R will sometimes be referred to as prior and post readiness respectively.

Step 2: Calculate student progress. Student progress is given simply by the difference in projected readiness from the end of 8th grade to the end of 9th grade.

Step 3: Construct student progress by component. In addition to overall student progress, we also estimate progress for each component as the difference between 9th grade readiness and predicted readiness after controlling for 8th grade student outcomes.

Step 4: Evaluate the student and school-level progress data with an equity lens. We measure student and school progress using models that distinguish between differences in progress within schools by student demographic status and between schools by composition by demographic subgroups.

Step 5: Apply shrinkage. To improve the accuracy of progress estimates, we apply a reliability adjustment estimation (i.e. statistical shrinkage) to the estimates of school progress. Therefore, all the reported progress estimates are reliability-adjusted metrics.

Table 9 presents a progress report for the student labeled as the "medium readiness student" in Table 3. Part 1 of the table (on the left side) provides information on prior 8th grade student outcomes and corresponding projected readiness (P). Part 2 of the table provides information for 9th grade student outcomes and the corresponding projected readiness (R). The data in part 2 exactly matches the data in the early warning report for this student in **Table 3**. Finally, the bottom part of the table presents the estimate of student progress (G) for this student.

Table 9. Student Progress Report for First-Year High School Student

Example Student / Medium Readiness Student					
Part 1			**Part 2**		
8th Grade Outcomes	**Prior Projected Readiness (P)**		**9th Grade Outcomes**	**Current Projected Readiness (R)**	
	Outcome	**Readiness**		**Outcome**	**Readiness**
Constant			*Constant*		1
Attendance	90	**7**	**Attendance**	92	**7**
Advanced Courses		**2**	**Advanced Courses**		**10**
Math	Yes		Math	Yes	
NA			Science	Yes	
Test Scores		**18**	**Test Scores**		**16**
ELA 8th Grade Score	297		Biology	67	
Math 8th Grade Score	295		Algebra 1	78	
Course Grades		**26**	**Course Grades**		**34**
Math	B		Math	A	
Science	B		Science	B	
English	C		English	C	
Social Studies	C		Social Studies	B	
Other Subjects	C		Other Subjects	C	
Total Readiness		**53**	**Total Readiness**		**68**
Readiness Status	50<R<75	**Q3: Medium**	**Readiness Status**	50<R<75	**Q3: Medium**

Progress Status	
Total Updated Readiness (R)	**68**
Total Prior Readiness (P)	**53**
Total Readiness Progress (G)	**15**

The example medium-readiness student in **Table 9** experiences very strong progress, with projected readiness increasing from 53 (about average readiness) to 68, an increase of 15 points. The increase in projected readiness is due to two components: enrollment in advanced courses in math and science in 9th grade, which yielded an increase in component scores of 8 points, and an improvement in course grades, which yielded an increase of 8 points. The grade in math improved from a B to an A and the grade in social studies improved from a C to a B. The other grades stayed the same. We can speculate that the improvement in the math grade for this student may have been affected by the fact that the student enrolled in advanced math in both years, but the student's teacher could draw on additional information to ascertain what supports helped this student improve.

As in the case of the school warning reports presented in previous section, the school accountability reports are provided for example Schools A (**Table 10A**) and B (**Table 10B**). The tables report progress and performance metrics for total readiness for all students and by subgroup. The progress and performance metrics have been estimated using reliability adjustment (shrinkage estimation) to increase the accuracy of the estimates.

School A is a high readiness school, with uniformly high prior readiness on average for all subgroups. Prior readiness values range from 80 to 83 on the 0/100 readiness scale for all subgroups, with an average prior readiness value for all students equal to 83. In contrast, School B is a low readiness school, with uniformly low prior readiness values ranging from 35 to 41 for all subgroups, with an average prior readiness value equal to 38 for all students.

The reports presented include three school progress/performance ratings:

- The school average student progress metric based on the 0/100 readiness scale (the focus of the high-level analyses in the previous section.
- A standardized school performance rating that facilitates rating schools on five different performance levels.
- The percentile measure of student progress and school performance.

The latter two metrics are designed to facilitate direct comparisons of progress across schools and allow classification of schools by performance levels.

The standardized school performance rating is statistically equivalent to the school progress metric used in the above analyses but is modified so that the metric can be reported on the same performance scale for all grades and is easy to interpret. The school performance rating linearly transforms the school progress metric to a scale like a z score, with a standard deviation equal to one but a mean equal to 3. If the estimated school effects are approximately normally distributed, 95% of the estimated ratings will lie in the interval 1 to 5 and 99% of the estimated effects will lie in the interval 0.5 to 5.5. This metric is used in Wisconsin school report cards applied to value-added estimates of school performance in math and ELA (Meyer and Christian, 2020). The school performance rating is reported in two ways: (1) two decimal accuracy (e.g., 4.38 for all students in School B) and (2) as a categorical performance level, where the rating is rounded to the nearest integer. The categorical rating is included, as in the early warning reports, to enable simple classifications of schools by five performance levels, P1 to P5:

School Performance Levels					
School Performance level	**P1: Very Low Performance**	**P2: Low Performance**	**P3: Medium Performance**	**P4: High Performance**	**P5: Very High Performance**
School Performance Rating	**0 - 1**	**2**	**3**	**4**	**5 - 6**

The school accountability report for School A shows that even though School A is a high prior readiness school, this school's average school progress is almost exactly equal to district average progress, -0.03 versus 0, and accordingly, the school performance percentile rank is 50%. The overall performance of school A is at medium level with a rating equal to P3 (on the 1 to 5 school performance scale). In contrast, although School B is a low prior readiness school, average school progress is substantially above the district average progress: school progress = 6.02 and percentile rank = 92%. Therefore, school B is a high-performance school with a performance rating equal to P4.

Table 10A. School Accountability Report: High Readiness School A

End-of-Year School Accountability Report on Readiness								
High School Year 1 (9th Grade)								
Comparability Control Variables: 8th Grade								
School A: High Readiness School								
			Prior Readiness and Progress			**Performance**		
Group variable	**Size**	**Sub-Group %**	**Prior Readiness (P)**	**District Progress Average**	**School Progress (G)**	**Rating Detail**	**Performance Rating**	**School Percentile**
All Students	1151	100%	82.98	0.00	-0.03	2.99	P3: Med	50%
Economically Disadvantaged:								
No	771	67%	82.87	0.82	0.30	3.07	P3: Med	53%
Yes	380	33%	83.21	-0.90	-0.82	2.81	P3: Med	43%
Race/Ethnicity:								
Other	1002	87%	83.35	1.59	0.35	3.08	P3: Med	53%
Black	76	7%	80.83	-0.50	-2.19	2.50	P2: Low	31%
Hispanic	73	6%	80.23	-1.14	-2.74	2.37	P2: Low	26%

Table 10B. School Accountability Report: Low Readiness School B

End-of-Year School Accountability Report on Readiness								
High School Year 1 (9th Grade)								
Comparability Control Variables: 8th Grade								
School B: Low Readiness School								
			Prior Readiness and Progress			**Performance**		
Group variable	**Size**	**Sub-Group %**	**Prior Readiness (P)**	**District Progress Average**	**School Progress (G)**	**Rating Detail**	**Performance Rating**	**School Percentile**
All Students	348	100%	38.47	0.00	6.02	4.38	P4: High	92%
Economically Disadvantaged:								
No	172	49%	39.99	0.82	7.39	4.70	P5: Very High	96%
Yes	176	51%	36.98	-0.90	5.09	4.17	P4: High	88%
Race/Ethnicity:								
Other	23	2%	34.95	1.59	8.09	4.86	P5: Very High	97%
Black	165	47%	40.81	-0.50	6.46	4.48	P4: High	93%
Hispanic	160	14%	36.56	-1.14	5.45	4.25	P4: High	89%

The economically disadvantaged, black, and Hispanic students have lower progress metrics compared to other students. For instance, in school A the average progress is negative for three subgroups: economically disadvantaged, black and Hispanic students. In school B while the progress is positive for all subgroups, the economically disadvantaged, black and Hispanic students have significantly lower progress. These differences across subgroups are sufficiently large that they should spur each school to investigate and address the underlying causes of these differences.

In summary, the progress and performance results presented in school reports confirm the value of evaluating student and school readiness using multiple student outcomes and value of assessing student and school performance with an equity lens. District reports for school progress and accountability reports, while not shown here, can be created in a similar format as district readiness reports shown in previous section. They would allow to see and compare the progress of each school as well as the overall district progress by demographic groups. The intent of the high-level district and school accountability reports is to spur reflection to address the challenges identified in the reports.

Analysis of Student Readiness and Progress by Group and School Composition

In addition to providing a within school analysis of performance across demographic groups already in the reports presented so far, EAAS also provides a comprehensive and equity focused analysis of the performance of demographic subgroups at the district level. The analysis aims to answer the equity-related question: *Are differences in student progress largely due to differences in individual student demographic characteristics or differences in school composition based on the proportions of students in demographic subgroups?* **Appendix C** describes the technical details of this analysis.

Tables 11 and **12** and **Figures 4** and **5** are designed to address this question. **Panel A** reports district average student readiness and progress for different levels (bins) of school composition with respect to economic disadvantage and race/ethnicity. The bins are broken down into intervals of 10 percentage points. The percent of students in a school in each bin is reported in **Column 2**. As an example, for school composition based on ED (**Table 11**), 5.56 percent of students are in schools with an ED proportion of 11 to 20%. The largest bin is for the ED proportion equal to 41 to 50% (21.39 percent of all students). For school composition based on race/ethnicity (**Table 12**), the largest bin is for the black or Hispanic student proportion equal to 91 to 100% (20.33 percent). **Columns 3** and **4** report average readiness for each subgroup and **Column 5** reports the difference in average readiness for the two subgroups. **Columns 6** to **8** report the same statistics for average student progress. Panel B in each table report the difference between average readiness and progress values for two composition levels (bins): 71 to 80% and 21 to 30%. These values complement the data in **Figures 4** and **5** in that they indicate whether increases in school composition with respect to economic disadvantage or race/ethnicity are associated with decreases or increases in average student readiness and progress.

As indicated in **Figures 4** and **5** (Chart A) and **Tables 11** and **12**, projected readiness differs strongly and negatively across schools as function of the school composition with respect to economic disadvantage and race/ethnicity, respectively. The effects of school composition are much larger than the effects of individual student status.[11]

[11] Note that the average difference in readiness for all students, -12.75 for ED status and -23.41 for race/ethnicity, are much larger than the difference in readiness for students in schools with the same composition, -5.56 for ED status and -10.46 for race ethnicity reported in the row labeled “Average Over Bins.” This result reflects the fact that readiness is much lower for both ED and Non-ED students in schools with high proportions of ED students and similarly for the race/ethnicity contrasts. The school composition differences in readiness are the largest factor in determining average readiness across subgroups for all students.

Figure 4. Average Student Readiness and Progress by ED Status and School Composition

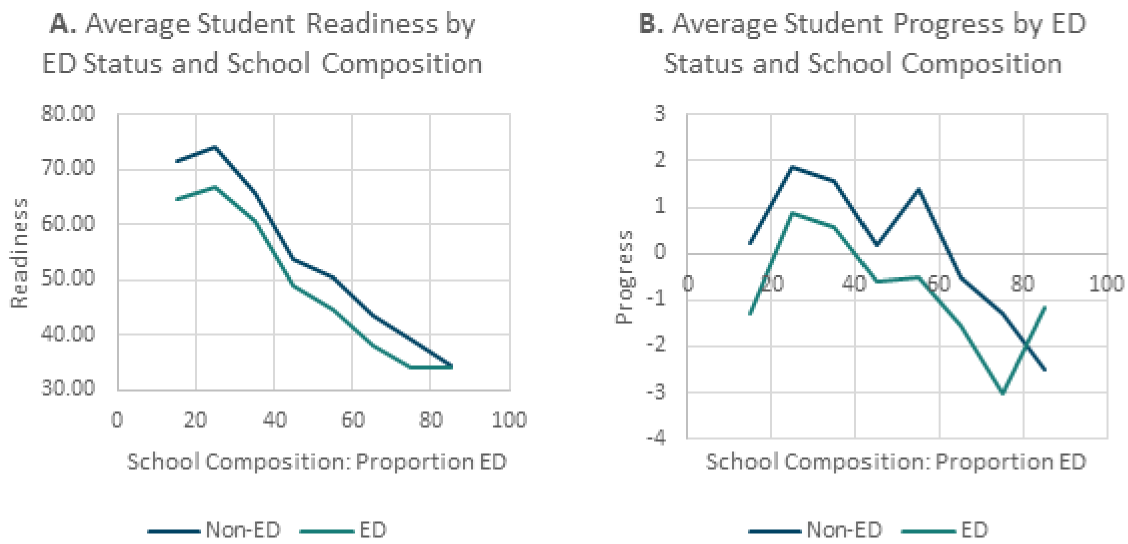
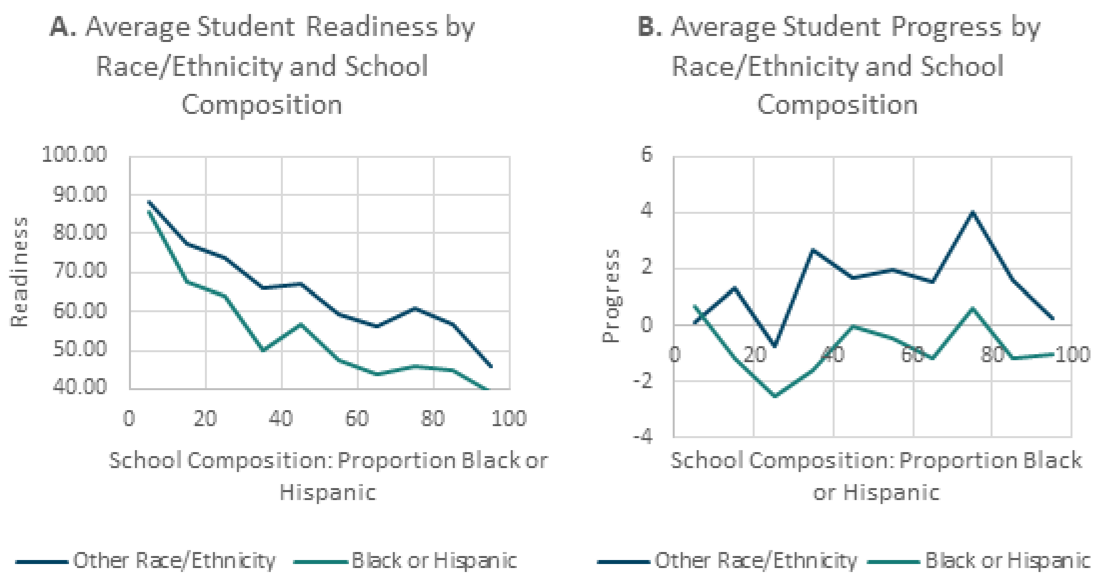


Figure 5. Average Student Readiness and Progress by ED Status and School Composition



The progress statistics reported in the **Tables 11** and **12** and in **Figures 4** and **5** (chart B) reveal that the differences in progress between the subgroups are relatively constant – the two average progress lines change as a function of school composition but are nearly parallel. This can also be interpreted that schools are not differentially effective with the different student subgroups included in analyses. The results also reveal that average progress decreases for schools with higher proportions of ED students for both Non-ED and ED students. The difference in average progress between school composition levels of 71-80% versus 21-30%

is -3.16 for non-ED students and -3.90 for ED students. In contrast, the average difference in progress for students in schools with the same composition is less, -1.21. For the race/ethnicity comparison, the opposite is true: average progress increases for schools with higher proportions of black or Hispanic students for both subgroups. **Panel B** of **Table 12** indicates that the difference in average progress between school composition levels of 71-80% versus 21-30% is 4.75 for Other students and 3.13 for black or Hispanic students. The average difference in progress for students in schools with the same composition is almost as large, but in the opposite direction, -2.28.

Table 11. Average Student Readiness and Progress by Group and School Composition: Economic Disadvantage

District Average							
School Composition	**Percent of Students in Bin**	**Student Readiness R**			**Student Progress G**		
		Economically Disadvantaged			**Economically Disadvantaged**		
(1)	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**	**(7)**	**(8)**
Panel A:	**-**	**No**	**Yes**	**Diff.**	**No**	**Yes**	**Diff.**
11 to 20%	5.56	71.61	64.48	-7.13	0.24	-1.30	-1.54
21 to 30%	12.63	74.22	66.74	-7.48	1.88	0.87	-1.01
31 to 40%	14.58	65.75	60.49	-5.26	1.57	0.58	-0.99
41 to 50%	21.39	53.79	49.19	-4.60	0.20	-0.58	-0.78
51 to 60%	20.81	50.38	44.58	-5.79	1.39	-0.53	-1.92
61 to 70%	15.37	43.54	38.13	-5.41	-0.51	-1.57	-1.06
71 to 80%	8.40	39.17	34.07	-5.10	-1.28	-3.04	-1.76
81 to 90%	1.00	34.61	33.99	-0.62	-2.49	-1.15	1.34
Average Over Bins	-	55.39	49.84	-5.56	0.60	-0.61	-1.21
All Students	100.00	58.94	46.19	-12.75	0.82	-0.90	-1.72
Panel B:							
71/80% – 21/30%	Diff:	-35.05	-32.67	2.38	-3.16	-3.90	-0.75

Table 12. Average Student Readiness and Progress by Group and School Composition: Race/Ethnicity

District Average							
School Composition	Percent in Bin	Student Readiness R Race/Ethnicity			Student Progress G Race/Ethnicity		
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Panel A:		Other	Black/ Hispanic	Diff.	Other	Black/ Hispanic	Diff.
0 to 10%	3.43	88.34	85.52	-2.83	0.11	0.70	0.59
11 to 20%	6.51	77.40	67.55	-9.84	1.34	-1.16	-2.50
21 to 30%	4.16	73.63	63.90	-9.73	-0.72	-2.53	-1.81
31 to 40%	13.19	66.04	50.23	-15.81	2.67	-1.58	-4.26
41 to 50%	8.94	66.87	56.90	-9.98	1.65	-0.01	-1.67
51 to 60%	6.04	59.23	47.40	-11.83	1.93	-0.47	-2.40
61 to 70%	8.51	56.24	43.79	-12.45	1.51	-1.14	-2.65
71 to 80%	8.03	60.88	46.05	-14.83	4.03	0.59	-3.43
81 to 90%	11.85	56.64	44.89	-11.76	1.60	-1.15	-2.75
91 to 100%	29.33	46.03	39.29	-6.75	0.26	-1.03	-1.29
Average Over Bins		59.29	48.84	-10.46	1.39	-0.88	-2.28
All Students	100.00	67.91	44.50	-23.41	1.59	-0.88	-2.47
Panel B:							
71/80% – 21/30%	Diff	-12.75	-17.85	-5.10	4.75	3.13	-1.62

For the ED comparison, we conclude that the difference in progress for Non-ED and ED students due to school composition is more than twice as important as the difference due to student ED status, although both differences are negative. For the race/ethnicity comparison, differences in progress for Black and Hispanic, and other students due to school composition are also large, but positive, and the differences due to race/ethnicity are uniformly negative, but variable. For both comparisons, differences due to school composition are large, but with opposite effects. This is an intriguing finding that deserves exploration to assess the determinants of the difference.

The analysis provided in this section looks deeper into the aspects of inequality within the district. It uses the results generated through EAAS to provide a larger picture of the performance of different demographic groups of students within the district. As such this analysis can be very relevant for reflection, accountability and decision making from equity perspective.

IV. TAKEAWAYS AND STEPS AHEAD

Traditional systems of accountability often have been deficit-oriented, have overly relied on test scores, and have reflected outcomes that are too late in a student's education experience to meaningfully improve. Yet when designed well, accountability can be a tool for equity—by highlighting where inequities are occurring and informing actionable strategies to combat them. The system described here brings together predictive analytics and early warning systems with accountability systems and the modifications of these measurement approaches that would transform accountability systems from threats to tools for improvement.

Including predictive analytics and early warning measures in accountability systems could identify opportunity gaps and root causes of inequities early enough for educators (and families) to act, so that more distal outcome gaps in high school graduation rates and even post-secondary well-being can be addressed while schools, students, and caregivers still have a chance to make real change. An improved and equity-focused accountability system could act less like a verdict on student and school performance and more like an actionable roadmap to improve students' life outcomes.

The research presented here demonstrates that such a system is feasible and produces valid results. Integrating predictive analytics and early warning systems into the traditional framework of accountability presents a promising avenue for creating an equity-aligned system—a system that is responsive to the schools and communities where the measures are being implemented. Future research is needed to build upon the promising findings here to further refine the linkage between early warning and accountability systems, develop tools for educators, parents, and other stakeholders to use and act upon the information provided, and expand the level of community engagement in developing such a system. Specifically, we recommend that future research build upon the promising findings here to:

- Work with parents and educators to develop tools they and other stakeholders would use to understand and act upon the information provided.
- Explore how the readiness index can be integrated with other accountability measures (e.g., what weight should it carry, does it make other metrics redundant, how does it change rankings of schools).
- Let users "test drive" the prospective metrics and user tools and understand the actions users are likely to take to influence them.
- Expand the system to use middle-grade and high school indicators for grades beyond 9 to predict workforce/post-secondary education outcomes, providing an even more direct link between schooling and post-secondary well-being.

We envision that the connections between the early warning and accountability branches of EAAS could be strengthened by incorporating goal setting and student learning objectives (SLOs) into the system. Student and school readiness goals (overall and by component) could be informed by historic evidence on student progress, thereby allowing students and educators to set goals that are ambitious *and* realistic.

Use Data to Support Students and Schools: Minimize Unintended Consequences

This study has focused on the feasibility of building an integrated system focused on prospective/early warning and retrospective/accountability analytics. EAAS is motivated by the assumption that this data can be used constructively to improve student and school outcomes. An important next step is to consider how to implement such a system to minimize possible unintended consequences and to engage with stakeholders and community members to iteratively build and refine the system.

We have addressed one such challenge: how to incorporate grades/GPA in the system without spurring grade inflation (see also **Appendix B**). Relative to traditional accountability systems which are based on a very limited set of outcomes (often only math and ELA test scores), EAAS has the advantage of including a broad and diverse set of student outcomes. Moreover, these outcomes are combined to create a composite measure of readiness. As a result, opportunities to game the system by focusing on a limited set of outcomes may be reduced.

Nonetheless, there is always the possibility that identifying students as having low or very low readiness could stigmatize them such that less support would be provided to these students. This response could realistically occur in an early warning system if there were insufficient safeguards in place. This response may be much less likely to occur in EAAS given that EAAS is designed to evaluate the progress of students at the end of the school year and over multiple school years. Indeed, the most comprehensive version of EAAS (see **Step 4** below) tracks and evaluates the effects of actions spurred (or not spurred) by prospective early warning metrics.

Implementation of EAAS: Options for Phased or Partial Implementation

We have demonstrated that the comprehensive version of EAAS presented in this report provides actionable, equity-aligned information on student and school readiness and progress for a broad set of student outcomes. However, no districts or states to our knowledge have implemented integrated early warning and student progress (accountability) systems that include a broad set of student outcomes, linked to medium or long-term post-high school outcomes. However, we believe that all districts and states have sufficient data to implement basic versions of EAAS that could build on, for example, their:

- Current accountability systems, by expanding their focus to include student outcomes beyond math and ELA tests.
- Experience with growth, value-added, and student growth percentile (SGP) models.
- Experience with systems that include features of early warning systems and/or goal setting and student learning objectives (SLOs).

Initial versions of EAAS could be based on available student data and then be expanded as additional data becomes available. We assume that all implementations of EAAS include its key components: prospective early warning metrics, retrospective progress-based accountability metrics, comprehensive equity-aligned analytics in both systems, and, as discussed above, data-informed goal setting. Below we consider one of the many possible strategies for phased implementation of a comprehensive version of EAAS. The data elements added in each phase are listed.

- **Step 1. Basic EAAS**
    - Math and ELA test scores
    - Student attendance and chronic absenteeism status
    - High school graduation projections not yet implemented. Composite measures of student readiness based on subjective weighting of outcomes.
    - Grades: 8 and 9 (the transition to high school grades)
- **Step 2. Expansion to broader set of high school outcomes**
    - Course enrollment and grade/GPA data
    - High school graduation status, college attendance, and college graduation. Calibrate models to be able to create composite readiness metrics
- **Step 3. Expand student and school outcomes and grades**
    - Possible outcomes: social emotional competencies, student wellness, culture and climate
    - Interim test scores
    - Career and technical education (CTE) outcomes
    - Grades: 3-12 and K-2, if possible
- **Step 4. Document actions to support students and schools and their impact**
    - Track actions, programs, and interventions
    - Evaluate impacts of actions, programs, and interventions
- **Step 5. Expand post-high school and related high school outcomes**
    - Post-high school CTE and workforce outcomes. Calibrate models to be able to create composite readiness metrics
- **Step 6 (and earlier). Evaluate the impact of EAAS**
    - Technical quality of the data, statistical models, and metrics
    - Effects on student outcomes and readiness for all types of students and schools?
    - Are their unintended consequences?
    - System redesign in collaboration with stakeholders and community members

In summary, successful implementation of EAAS will inevitably require constant formative evaluation of the system and redesign of the system informed by that evaluation.

APPENDIX

Appendix A

Statistical Models and Model Estimates of Early Warning System Models of Projected Readiness

This appendix describes the statistical model, formulas, and statistical results used to construct measures of projected student readiness *R* and associated high school graduation probabilities.

High School Graduation Status Model

The dependent variable to be predicted is the medium-term outcome: high school graduation status, given by:

High school graduation status = D =

0. Did not graduate within four years
1. Non-honors diploma
2. Honors diploma

Given that this dependent variable is a discrete, multi-valued, and ordered variable, the best statistical model to realistically represent this outcome is an ordered probit or logit model (Maddala, 1983; Daykin & Moffatt, 2002; Greene, 2017; Wooldridge, 2010). Although both models yield nearly identical results, we use the ordered probit model because it is most compatible with the regression models used elsewhere in this report.

The ordered probit model is defined by the probabilities associated with each of the three high school graduation outcomes. These probabilities are a function of: (1) an equation that takes the same form as a linear regression model, but with the latent variable U_{ik} as the dependent variable, (2) threshold parameters to determine assignment to each of the three outcomes, and (3) the link function that is used to calculate the probabilities. In probit models the link function is the standard normal distribution function $\Phi[.]$.

The latent variable equation is given by:

$$U_{ik} = W_{ik}\eta^* + e_{ik}^*$$

where:

- The vector W_{ik} for student i in school *k* represents all high school outcomes/predictors included in the model
- η^* represents the corresponding coefficient vector on the latent outcome scale

- e_{ik}^* represents a random error term assumed, given the assumptions of the probit model, to be normally distributed with a normalized standard deviation equal to $\sigma^* = 1$.

Let c_{12}^* and c_{23}^* represent the threshold parameters that determine the boundaries between graduation outcome levels: (a) 1 and 2 and (b) 2 and 3, respectively. The probability that a student earns an honors diploma is given by:

$$P_{3ik} = \Phi\left[\frac{W_{ik}\eta^* - c_{23}^*}{\sigma^*}\right]$$

Similarly, the probability that a student earns either diploma (non-honors or honors) is given by:

$$P_{2\&3ik} = \Phi\left[\frac{W_{ik}\eta^* - c_{12}^*}{\sigma^*}\right]$$

The model parameters -- $\eta^*, c_{12}^*, c_{23}^*$ -- can be estimated (calibrated) using any of the widely available software programs for estimating ordered probit models using historic longitudinal data on high school graduation status and outcomes/predictors in high school year 1. We estimated the models using the R procedure "polr."[12]

Projections of readiness on the latent scale defined by the ordered probit model are given by:

$$R_{ik}^* = W_{ik}\hat{\eta}^*$$

where the star superscript denotes the latent scale and $\hat{\eta}^*$ denotes the estimated coefficient vector on the latent scale parameters (with the ^ symbol added to denote an estimated parameter). As indicated in the Table A1 below, the sample size is sufficiently large such that the coefficients are estimated with very high precision. We linearly transform the calibrated coefficients so that resulting readiness projections R_{ik} are measured on a 0/100 scale. Since in a large district or statewide data set there are inevitably students who have very low and very high projected readiness values, we anchor the 0/100 scale on student readiness values at the 1st and 99th percentile of projected readiness values. This ensures that students have a realistic chance to earn scores at the bottom and (especially) top of the 0/100 range. Readiness projections for each of the four components included in the model (attendance, advanced course taking, test scores, and grades/GPAs) are computed using the same calibrated coefficients.

[12] See polr: Ordered Logistic or Probit Regression in MASS: Support Functions and Datasets for Venables and Ripley's MASS (rdrr.io).

Table A1 reports the calibrated coefficients for the high school graduation model transformed to the 0/100 readiness scale. **Table A2** shows the how projected readiness values translate into probabilities of graduating from high school with non-honors and honors diplomas.

Table A1. Calibrated Coefficients of High School Graduation Model

	Variable	Est	SE	t-value
	(Intercept)	1.15		
Attendance	Attend 81-90%	5.07	0.36	14.15
	Attend 91-93%	6.97	0.40	17.52
	Attend 94-96%	8.39	0.38	22.22
	Attend 97-98%	9.18	0.39	23.28
	Attend 99-100%	11.51	0.40	28.84
Adv. Courses	Math	5.39	0.28	19.58
	Science	4.82	0.33	14.68
Test Scores	Test 50-59	1.42	0.21	6.67
	Test 60-64	2.92	0.24	12.27
	Test 65-75	6.12	0.16	37.45
	Test 76-85	9.57	0.21	46.11
	Test 86-100	12.18	0.25	48.90
Course Grades/GPAs	GPA Math	3.18	0.12	27.00
	GPA Science	1.48	0.13	11.61
	GPA English	2.72	0.13	20.81
	GPA Social Studies	2.29	0.13	17.70
	GPA Other Subjects	2.12	0.12	16.99
	0/100 Scale multiplier	15.30		
	Pseudo R squared	0.73		
	N	> 10,000		

Table A2. Graduation Probabilities for Selected Readiness Thresholds for Classifying Students by Readiness Level

Graduation Probabilities				
	Threshold	Graduation Probability		
Student Readiness Level	Projected Readiness R	Non-Graduation	Non-Honors and Honors Diploma	Honors Diploma Only
Q1: Very Low Readiness: R Index <= 25	0	92%	8%	0%
	25	38%	62%	0%
Q2: Low Readiness: R Index = (26,50)	26	38%	62%	0%
	50	3%	97%	6%
Q3: Medium Readiness: R Index = (51,75)	51	3%	97%	7%
	75	0%	100%	54%
Q4: Very High Readiness: R Index > 75	76	0%	100%	57%
	100	0%	100%	96%

Appendix B

The Statistical Model Adjusting Course Grades to Eliminate a Grade Inflation Incentive

The GPA variables in high school year 1 (9^{th} grade) for each subject area are adjusted to eliminate any incentive for schools to artificially inflate grades. This approach allows grades to be included in an integrated early warning and accountability system, thereby retaining the strong predictive power of grades in early warning metrics. It is important to note that in EAAS schools are free to adopt any grading policy and may give lower or higher grades to students than other schools. Indeed, the student reports convey the actual grades earned by students (not adjusted grades). The grade adjustment is applied when reporting projected readiness values.

The method for adjusting grades is very similar to simply subtracting from student grades in each subject the average school-level grade in that subject for students in that high school grade/year. However, we recognize that schools on average may assign higher or lower grades based on students' prior academic preparation in 8^{th} grade and all prior student outcomes in 8^{th} grade that predict grades in 9^{th} grade. In recognition of this fact, we subtract from student grades the average school-level grade after controlling statistically for the same set of 8^{th} grade predictors included in the 9^{th} grade progress model discussed in the main text and in Appendix C. This grade adjustment is applied separately to grades in each subject area. We obtain more accurate estimates of both student readiness and school-level progress using this approach. Below, we summarize the performance of the grade adjustment models.

The explanatory power of these models, measured by the R-square statistic, ranges from 0.45 to 0.53, with the lowest R-square values in the models of science and other subject grades. The standard deviation of school effects measures the degree to which schools assign different grades on average for students with the same prior academic preparation and all prior student outcomes in 8^{th} grade. This standard deviation ranges from 0.31 to 0.36 for all subjects other than science. The standard deviation of effects for science grades is slightly higher: 0.47. These results indicate that school grading practices differ somewhat among schools even after controlling for students' prior 8^{th} grade outcomes. For example, the difference in average grades in math between comparable students in schools that assign grades one standard deviation higher versus one standard deviation lower than the average district school equals 0.72, a grade different about equal to the difference between a B- and C grade.

Appendix C

Statistical Models of Student and School Progress on Projected Student Readiness

This appendix describes the statistical models of student progress included in EAAS, associated formulas, and the statistical results used and reported in this study.

The models presented below actually serve two technical functions. One, the progress models, like traditional growth or value-added models, aim to estimate student and school-level differences in student outcomes and projected readiness measured at the end of 9th grade, controlling for student outcomes measured at the end of 8th grade. Two, the progress models extend the capacity to project student readiness from 9th grade to 8th grade since the progress model can also be used to construct prospective student readiness projections using data measured at the end of 8th grade.[13] The section in the report on early warning systems focused on student and school reports for 9th grade, but the progress reports necessarily report projections for both 8th and 9th grade, the prior and post years.

Whereas growth models typically focus on single student outcomes such as math and English language arts (ELA) achievement, the dependent variables in this report – projected readiness or readiness components – are composite variables comprised of the multiple high school outcomes that are included as predictors in the high school graduation status model. Similarly, whereas growth models typically include a limited set of student outcomes as predictors or control variables (for example, prior test scores in math and/or ELA), the EAAS progress model includes a set of 8th grade variables that fully match the four components included in the high school graduation status model. The progress model is a retrospective model; it is estimated annually using up-to-date post-year and pre-year data and yields contemporaneous estimates of student progress and updated estimates of projected readiness using the prior 8th grade data. The resulting reports provide information on student progress from the end of 8th grade to the end of 9th grade and naturally incorporate student outcome and projected readiness data from both years. The difference in projected readiness equals the student-level measure of progress. Thus, the reports provided at the end of the school potentially provide students, parents, and educators with the data, supplemented by their own experiences, to diagnose and reflect on challenges and opportunities for improvement.

[13] Alternatively, the high school graduation model can be estimated using 8th grade predictors and projected readiness measures can be constructed given 8th grade predictors. The computed 8th grade readiness measures were very similar for the two approaches.

The EAAS progress models incorporate key statistical features of commonly used growth-type models, including value-added and growth models[14] and student (and mean) growth percentile models (SGP/MGP)[15] (see citations in footnotes). We use the term "progress model" to signal that the EAAS progress models differ in some important respects from the commonly used growth models. EAAS models substantially expand these models to highlight equity. In addition, they are also based on composite variables that combine multiple student outcomes rather than single test score variables and they include as control variables prior grade/year measures of these variables. **EAAS Progress Model 1**, presented below, most closely resembles existing growth and value-added models.

The Progress Models

The high school graduation model addresses the fact that the dependent variable is an ordered, discrete outcome by employing an ordered probit model (as discussed in Appendix A). The calibrated coefficients from this model, combined with annually updated 9th grade outcomes, produce projected readiness values R_{ik} and components of readiness $R_{ik(m)}$ for component *m* for student *i* in school *k* at the end of 9th grade. As a result, it is appropriate to use linear regression and multilevel regression models to model projected readiness and produce measures of student and school average progress.

Four equity-aligned progress models and metrics are used to construct EAAS metrics.

- Model 1/Level 1: The Average Progress Model
- Model 2/Level 2: The Average Progress Model with Differential Student Progress
- Model 3/Level 3: The Differential Student and School Progress Model
- Model 4/Level 4: The Generalized Differential Progress Model: Systemic Differences Between and Within Schools

Although all models share the same basic structure, each successive model provides deeper levels of equity-aligned information on student and school progress. Below we present the Average Progress Model. We then describe how this model is extended to obtain the other three equity-aligned models.

Progress Model 1: The Average Progress Model. The average progress model consists of two parts. The first part provides estimates of predictions of student readiness (P_{ik}) and student progress (G_{ik}) for each student, where the subscripts *i* and *k* index students and schools in 9th

[14] Value-added and growth model references: Willms and Raudenbush (1989), Sanders and Horn (1994), Meyer (1997), McCaffrey et al (2004), Kane, McCaffrey, Miller and Staiger (2013), Chetty, Friedman and Rockoff (2014), Meyer and Dokumaci (2105), Guarino, Reckase and Wooldridge (2015), Koedel, Mihaly and Rockoff (2015), and Gawade and Meyer (2016).
[15] SGP model references: Betenbenner (2009), Guarino et al (2015), and Lockwood and Castellano (2015).

grade, respectively. The estimate of student progress (G_{ik}) is subsequently used as the dependent variable in all four models of student progress. The first part of Model 1 is given by:

$$R_{ik} = \xi + W_{8ik}\lambda + G_{ik} \quad (1)$$

where W_{ik}= the vector 8th grade outcomes as predictors, λ is the corresponding coefficient vector,ξ is the model intercept, and, as indicated above, G_{ik}= student progress after controlling for prior 8th grade outcomes, normalized to have mean zero (since it acts as an error component). The 8th grade outcomes included in the model were discussed in the text. As discussed above, one of the important features of the model is that predicted readiness, given by:

$$P_{ik} = \xi + W_{8ik}\lambda \quad (2)$$

is an updated estimate of projected readiness based on 8th grade outcomes. Hence, student progress is simply given by the difference in projected readiness from the end of 8th grade to the end of 9th grade:

$$G_{ik} = R_{ik} - P_{ik} \quad (3)$$

The progress tables reported in the text exploit this fact and report all student information needed to construct the estimate of student progress, namely, student outcome data from 8th and 9th grade, the calculated projected readiness scores for each component and total readiness in each year.

The model is estimated with 9th grade school effects included in the model so that the estimated slope coefficients are estimated using only within school variation in both the dependent and predictor variables. Including school effects ensures that there is sharp separation between student and school contributions to readiness. As indicated in (2), predicted readiness P_{ik} does not include the school effect; it is included in the progress measure G_{ik} since this measure intentionally includes both between-school and within-school student progress.

Similar models are also estimated for each of the four readiness components with the same set of predictors included in the model. These models yield separate coefficient and student progress estimates for each component (subscripted by component index *m*): $\xi_m, \lambda_m, P_{mik}, G_{mik}$. The parameters of the component models sum exactly to the corresponding parameters from the model of total readiness. The separate component estimates of student progress are useful because they provide data on which components are the source of low versus high overall student progress and thus can potentially inform diagnoses of the effectiveness of different actions spurred by early warning data provided at the beginning of the school year.

Model discussion. As discussed above, the EAAS progress models differ from common growth-type models in that the primary dependent variable, total projected readiness R_{ik}, is a composite measure that is composed of a broad set of student outcomes. Similarly, the models include a broad set of predictors that control for prior student differences. The outcomes in the separate component models are less broad, by design, but they also control for a broad set of predictors. As indicated below, this model design produces models with very strong predictive power, as measured by the R-square statistic. As a result, the EAAS progress models, compared to common growth models that focus on achievement in single subjects and include more limited predictors, may be much less prone to omitted variable bias.

One possible approach to further ensuring that bias is limited is to add additional student control variables to the model; in particular, student demographic variables such as economic disadvantaged status and race/ethnicity and/or school-level means of these variables. This option has generally not been permitted in models used as a part of federal (ESSA and NCLB) required accountability systems. SGP models and the layered value-added model of Sanders and Horn (1994) also do not include these variables but instead include multiple lags of prior student predictors. We have deliberately excluded demographic variables from the model as *control* variables because a primary focus of EAAS is to identify differences in progress of students with different demographic characteristics, not control away these differences.[16]

The student progress measures obtained from **Model 1** are used as the dependent variables in all four progress models considered in this report. The primary focus of all four models is to discover how student progress is affected, both within and across schools, by student and school demographic factors. **Model 1**, the Average Progress Model, includes fixed school effects in the model. Average progress at the school level is given simply by the average of the student level progress values for each school (hence, the name of the model). This process is represented formally by the following analysis of variance/variance components model:

$$G_{ik} = \alpha_k + u_{ik} \tag{4}$$

where α_k = the average progress school effect and u_{ik} = the student error component (residual). Estimates of student and school average progress based on Model 1 are provided in the report. To convey school-level progress data in a format that highlights school differences in performance, it is useful to report the average school progress effects on a standardized performance scale that is easy to interpret. We adopt the reporting scale that has been successfully used in Wisconsin applied to value-added estimates of school performance in math and ELA (Meyer and Christian, 2020;

[16] We are especially reluctant to include as control variables school-level means of either prior student outcomes or demographic variables in the progress model. Growth models that include school mean variables essentially eliminate the question of whether low readiness- students are systematically enrolled in schools with lower or higher average progress because the models explicitly control for school composition. Willms and Raudenbush (1989) and Meyer (1997).

WI_DPI_School_VA_Technical_Report_2020.pdf). The school progress performance scale transforms the school progress effect to a scale like a z score, with a standard deviation equal to one but a mean equal to 3. If the estimated school effects are approximately normally distributed, 95% of the estimated effects will lie in the interval 1 to 5 and 99% of the estimated effects will lie in the interval 0.5 to 5.5.

***Other progress models*. Progress Models 2**, **3**, and **4** expand **Model 1** to provide more detailed information on student and school average progress by student and school demographic subgroups and by school composition. To measure the full (rather than partial) differences between different subgroups, the models are estimated separately for each demographic variable. The models thus measure descriptive differences in progress between subgroups rather than the causal effects of subgroup status.[17] We consider two demographic group variables: economic disadvantage (ED) and race/ethnicity. In the ED model, differential progress effects are included for non-ED (subgroup 0) and ED (subgroup 1) students. In the race/ethnicity model, differential progress effects are included for black, Hispanic, and Other race/ethnicity-group students. The four progress models are presented in the following table.

Table C1. Progress Model Descriptions

#	Model	School Effect for: (a) All Students (b) By Student Subgroup (c) By Student Subgroup and School Composition by Subgroup	Subgroup Effect
1	The Average Progress Model	All	None
2	The Average Progress Model with Differential Student Progress	All	Yes
3	The Differential Student and School Progress Model	Student Subgroup	Yes
4	The Generalized Differential Progress Model: Systemic Differences Between and Within Schools	Student Subgroup and School Composition	Yes

[17] An alternative approach is to include multiple demographic variables. This approach maximizes the predictive power of the model but obscures differences in progress between subgroups because the multiple demographic variables are typically highly correlated. We have deliberately excluded multiple demographic variables from the model as *control* variables because a primary focus of EAAS is to identify differences in progress of students with different demographic characteristics, not control away these differences.

Model 2 expands the progress model to include differential student progress effects by subgroup (but not differential school progress effects). As in **Model 1**, α_k = the school effect is restricted to being the same for all subgroups (i.e., all students).

Model 3 expands the progress model to allow differences across schools in average progress above and beyond the differences in student progress included in **Model 2**. Thus, the model provides separate estimates of school progress effects for each demographic sub-group.[18] The average school progress effect for a given school from **Model 2** is simply the weighted average of the two separate progress effects in **Model 3**. Note that the centered effects can be highly, even perfectly correlated.[19]

Model 4 generalizes **Model 3** by both allowing school progress effects to differ for different demographic subgroups (as in **Model 3**) and allowing these school effects to differ systematically with the school composition of these subgroups. In other words, **Model 4** considers both within and between school differences in school progress for different demographic subgroups.

One of the practical challenges of reporting and making decisions based on the differential school progress estimates is that the precision of those estimates is strongly affected by the number of students in each subgroup. In schools with modest student populations, the number of students in one subgroup or another (for example, different race/ethnicity sub-groups) could be very small. Differential progress estimates are essentially unknowable for subgroups in schools with small sample sizes. Note, however, that progress effects for the subgroups with adequate sample sizes can be meaningfully compared with comparable estimates from other schools and with estimates based on the entire district. The problem of reporting metrics for subgroups with small sample sizes is well known. We implement a solution to this problem that has been demonstrated to be feasible and effective when applied in several districts and states. We improve the accuracy of the school effect estimates by applying reliability adjustment (shrinkage estimation) methods. In **Models 3** and **4**, shrinkage estimation borrows information from all estimated differential school effect estimates in each school. Thus, if school-level progress tends to be correlated (generally positive correlated) across different subgroups in the same school, the precision of all estimates can be improved. Subgroup estimates based on very limited sample sizes are very noisy and thus borrow much information from the estimates based on larger sample sizes. Estimates based on large sample sizes are not much changed by application of shrinkage. Shrinkage estimation formulas are provided in Meyer and Pier (2018), Meyer and Christian (2020), and in the references listed in a footnote in the discussion of **Model 1**.

---

[18] Wisconsin reports a similar differential effects value-added measure in the state school report card (Meyer and Christian, 2020; (WI_DPI_School_VA_Technical_Report_2020.pdf).

[19] This implies that the differences in subgroup effects for all school will be very similar unless the variances of the centered effects are quite different,

Table C2 reports estimates of **Model 1** for total readiness. Separate models were also estimated for the four readiness components: attendance, advanced courses, test scores, and grades/GPA.

Table C2. Estimated Coefficients of Model 1: Average School Progress Model

Grade 8 Calibrated Coefficients				
	Variable	**Est**	**SE**	**t-value**
Attendance	Attendance 81 to 90	6.82	0.27	25.64
	Attendance 91 to 93	11.30	0.28	41.10
	Attendance 94 to 96	13.72	0.27	50.85
	Attendance 97 to 98	16.18	0.28	58.68
	Attendance 99 to 100	18.07	0.29	61.27
	High Math Course	1.60	0.43	3.70
Course GPAs	Course GPA in High Math	2.85	0.12	23.03
	Course GPA in Low Math	2.36	0.08	29.45
	Course GPA in Science	2.39	0.08	30.76
	Course GPA in English	1.70	0.08	21.44
	Course GPA in Soc. Science	1.95	0.08	24.67
	Course GPA in Other Subjects	1.42	0.08	18.55
Grade 8 Test Indicators, Scores, and Interactions	Grade 8 ELA State Test Score	0.09	0.00	36.88
	Take Math End-of-Course Exam	18.94	1.31	14.45
	Take Both Grade 8 and Math End-of-Course Exam	-1.05	2.23	-0.47
	Math Grade 8 Test Score if Only Exam Taken	0.15	0.00	57.03
	Math End-of-Course Test Score if Only Exam Taken	0.50	0.02	31.45
	Math Grade 8 Test Score if Both Math Exams Taken	0.11	0.01	11.81
	Math End-of-Course Test Score if Both Exams Taken	0.27	0.03	10.52
Model Statistics	R squared	0.73		
	Standard Deviation of Dependent Variable	25.27		
	Standard Deviation of Student Error	11.30		
	Sample Size N	>10k		
	Fixed School Effects	Yes		
	Standard Deviation of School Effect	4.55		
	Intercept (Weighted Average of School Effect)	-64.78		

REFERENCES

Adams, C. M., Ford, T. G., Forsyth, P. B., Ware, J. K., Olsen, J. J., Lepine Sr, J. A., ... & Mwavita, M. (2017). Next Generation Accountability: A Vision for School Improvement under ESSA. *Learning Policy Institute*.

Alliance for Excellent Education. (2018). Ensuring Every Student Matters: What Is N-Size and Why Is It Important? *Alliance for Excellent Education.*

Beach, P., Thier, M., Lench, S. C., & Coleman, M. (2015). From Accountability to Actionability: Making Sense of Multiple Measures in Local Control Accountability Plans. Policy Brief. *Educational Policy Improvement Center*.

Betebenner, D. (2009). Norm-and criterion-referenced student growth. *Educational Measurement: Issues and Practice*, *28*(4), 42-51.

Bush-Mecenas, S., de Oca, D. M., Marsh, J., & Hough, H. (2018). "Seeing the whole elephant": Changing mindsets and empowering stakeholders to meaningfully manage accountability and improvement. *Education Policy Analysis Archives*, *26*, 10-10.

Carl, B., Richardson, J. T., Cheng, E., Kim, H., & Meyer, R. H. (2013). Theory and application of early warning systems for high school and beyond. Journal of Education for Students Placed at Risk (JESPAR), 18(1), 29-49.

California Department of Education. (2023, March 06). Growth Model.

Chetty, R., Friedman, J. N., & Rockoff, J. E. (2014). Measuring the impacts of teachers I: Evaluating bias in teacher value-added estimates. *American economic review*, *104*(9), 2593-2632.

Christian, M, & Meyer, R. H., (2020). Technical Report on the Wisconsin Value-Added Model: Target Group Update, Academic year 2018-2019. Education Analytics. https://dpi.wi.gov/sites/default/files/imce/accountability/pdf/WI_DPI_School_VA_Technical_Report_2020.pdf

Darling-Hammond, L., Cardichon, J., & Melnick, H. (2017). Identifying Schools for Support and Intervention: Using Decision Rules to Support Accountability and Improvement under ESSA. Policy Brief. *Learning Policy Institute*.

Daykin, A. R., & Moffatt, P. G. (2002). Analyzing ordered responses: A review of the ordered probit model. *Understanding Statistics: Statistical Issues in Psychology, Education, and the Social Sciences*, *1*(3), 157-166.

De Brey, C., Musu, L., McFarland, J., Wilkinson-Flicker, S., Diliberti, M., Zhang, A., ... & Wang, X. (2019). Status and Trends in the Education of Racial and Ethnic Groups 2018. NCES 2019-038. *National Center for Education Statistics*.

Edley, C., Jr., Koenig, J., Nielsen, N., & Citro, C. (Eds.). (2019). *Monitoring educational equity.* Washington, DC: National Academies Press. Retrieved from https://eric.ed.gov/?id=ED595713.

Frazelle, S. & Nagel, A. (2015). A Practitioner's Guide to Implementing Early Warning Systems. REL 2015–056. *Regional Educational Laboratory Northwest.* Retrieved from http://ies.ed.gov/ncee/edlabs.

Fox, J. H., & Balfanz, R. (2020). Keeping Students on Track to Postsecondary Success: Learnings from the Pathways to Adult Success (PAS) Initiative. *Teachers College Record*, *122*(14), 1-24.

Gawade, N. G., & Meyer, R. H. (2016). Measuring teacher effectiveness using value-added models of high school achievement. *Teachers College Record*, *118*(13), 1-32.

Greene, G. (2017). *Econometric Analysis*, 8th edition. Pearson Education Limited.

Guarino, C. M., Reckase, M. D., Stacy, B. W., & Wooldridge, J. M. (2015). A Comparison of Student Growth Percentile and Value-Added Models of Teacher Performance. *Statistics & Public Policy*, 2(1), 1-11.

Guarino, C. M., Reckase, M. D., & Wooldridge, J. M. (2015). Can value-added measures of teacher performance be trusted?. *Education Finance and Policy*, *10*(1), 117-156.

Harman, W. G., Boden, C., Karpenski, J., & Muchowicz, N. (2016). No child left behind: A postmortem for Illinois. *Education Policy Analysis Archives*, *24*, 48-48.

Hemelt, S. W., & Jacob, B. A. (2020). How does an accountability program that targets achievement gaps affect student performance? *Education Finance and Policy*, *15*(1), 45-74.

Hough, H., & Witte, J. (2016). Making Students Visible: Comparing Different Student Subgroup Sizes for Accountability. Policy Memo 16-2. *Policy Analysis for California Education, PACE*.

Johnson, E., & Semmelroth, C. (2010). The predictive validity of the early warning system tool. *NASSP Bulletin*, *94*(2), 120-134.

Kane, T. J., McCaffrey, D. F., Miller, T., & Staiger, D. O. (2013). Have we identified effective teachers? Validating measures of effective teaching using random assignment. *Research Paper. MET Project. Bill & Melinda Gates Foundation*.

Kane, T. J., Staiger, D. O., Grissmer, D., & Ladd, H. F. (2002). Volatility in school test scores: Implications for test-based accountability systems. *Brookings papers on education policy*, (5), 235-283.

Koedel, C., Mihaly, K. & Rockoff, J. E. (2015). Value-added modeling: A review. *Economics of Education Review*, *47*, 180-195.

Lee, J., Kim, N., Cobanoglu, A., & O'Connor, M. (2019). Moving to Educational Accountability System 2.0: Socioemotional Learning Standards and Protective Environments for Whole Children. Education Policy. *Nelson A. Rockefeller Institute of Government*.

Lockwood, J. R., & Castellano, K. E. (2015). Alternative statistical frameworks for student growth percentile estimation. *Statistics and Public Policy*, *2*(1), 1-9.

Longford, N.T. (1999). Multivariate shrinkage estimation of small area means and proportions. *Journal of the Royal Statistical Society*, *162*(Part 2), 227-245.

Longford, N. T. (2005). *Missing data and small-area estimation: Modern analytical equipment for the survey statistician*. Springer Science & Business Media.

Mac Iver, M. A., Stein, M. L., Davis, M. H., Balfanz, R. W., & Fox, J. H. (2019). An efficacy study of a ninth-grade early warning indicator intervention. *Journal of Research on Educational Effectiveness*, *12*(3), 363-390.

Maddala, G. S. (1983). *Limited-Dependent and Qualitative Variables in Economics*. Cambridge University Press.

McMahon, B. M., & Sembiante, S. F. (2020). Re-envisioning the purpose of early warning systems: Shifting the mindset from student identification to meaningful prediction and intervention. *Review of Education*, *8*(1), 266-301.
McCaffrey, D. F., Lockwood, J. R., Koretz, D., Louis, T. A., & Hamilton, L. (2004). Models for value-added modeling of teacher effects. *Journal of educational and behavioral statistics*, *29*(1), 67-101.
Meyer, R. H. (1997). Value-added indicators of school performance: A primer. *Economics of education Review*, *16*(3), 283-301.
Meyer, R. H., & Dokumaci, E. (2010). Value-added models and the next generation of assessments. *ETS Center for K–12 Assessment and Performance Management. Recuperado em*, *20*.
Meyer, R.H. & Pier, L. (2018). Small Numbers of Students Research, Report to the New York State Education Department.
Meyer, R.H. & Yu, Sara. (2023). Learning Lag and Recovery after COVID, Paper presented at the Annual Conference of the Society for Research on Educational Effectiveness, September 29. 2023.
Perdomo, J. C., Britton, T., Hardt, M., & Abebe, R. (2023). Difficult Lessons on Social Prediction from Wisconsin Public Schools. *arXiv preprint arXiv:2304.06205*.
Pierson, A., Frazelle, S., & Mazzeo, C. (2020). Adapting and Using Early Warning Indicators in Different Contexts. *Teachers College Record*, *122*(14), 1-24.
Rao, J.N.K. (2003). *Small Area Estimation,* New York, NY: Wiley-Interscience.
Raudenbush, S. W., & Bryk, A. S. (2002). *Hierarchical linear models: Applications and data analysis methods*, second edition. sage.
Sanders, W. L., & Horn, S. P. (1994). System (TVAAS): Mixed-Model Methodology. *Journal of Personnel Evaluation in education*, *8*, 299-311.
Seastrom, M. (2017). Best Practices for Determining Subgroup Size in Accountability Systems While Protecting Personally Identifiable Student Information. Institute of Education Sciences Congressionally Mandated Report. IES 2017-147. *Institute of Education Sciences*.
Snijders, T. A. (2011). Statistical models for social networks. *Annual review of sociology*, *37*, 131-153.
Spurrier, A., Aldeman, C., Schiess, J. O. N., & Rotherham, A. J. (2020). The Impact of Standards-Based Accountability. Making Next Year Count: Summer 2020 Accountability Series. *Bellwether Education Partners*.
TNTP (2016). Accountability under ESSA: How States Can Design Systems That Advance Equity and Opportunity. Working Paper. TNTP.
Torres, R. (2021), "Does test-based school accountability have an impact on student achievement and equity in education?: A panel approach using PISA", OECD Education Working Papers, No. 250, OECD Publishing, Paris, https://doi.org/10.1787/0798600f-en.
Wentworth, L., & Nagaoka, J. (2020). Early warning indicators in education: innovations, uses, and optimal conditions for effectiveness. *Teachers College Record*, *122*(14), 1-22.

Willms, J. D., & Raudenbush, S. W. (1989). A longitudinal hierarchical linear model for estimating school effects and their stability. *Journal of educational measurement*, *26*(3), 209-232.

Wooldridge, J. M. (2010). *Econometric analysis of cross section and panel data, second edition*. MIT press.