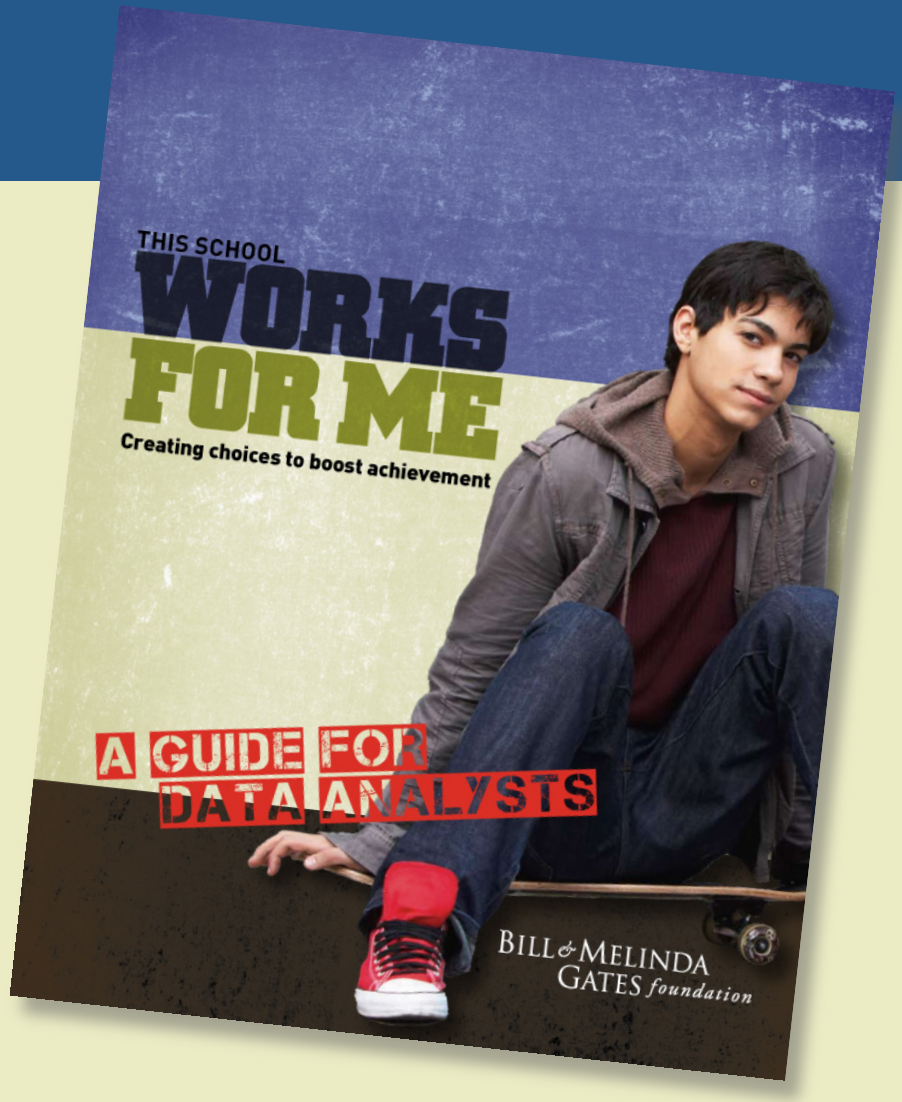
BILL & MELINDA GATES foundation

Supplement to

This School Works for Me:

Creating Choices to Boost Achievement

A Guide for Data Analysts



January 2010

INITIAL STEPS
Creating a Project Timeline

	Month 1	Month 2	Month 3	Month 4	Month 5+
Sample Timeline*					
Launch Project					
Identify district and/or external partner working team and develop project plan					
Compile district performance overview					
Conduct Analysis					
Gather necessary data					
Complete analysis					
Review results and iterate as necessary					
Engage Stakeholders					
Communicate purpose of analysis and seek support					
Communicate findings and ongoing strategy					
Implement recommendations (detailed in Portfolio Development Guide)					Ongoing implementation

Kick Off Meeting

Update Meeting

The whole process—from analysis to implementation—takes significant time and effort to complete so expectations need to be managed accordingly.

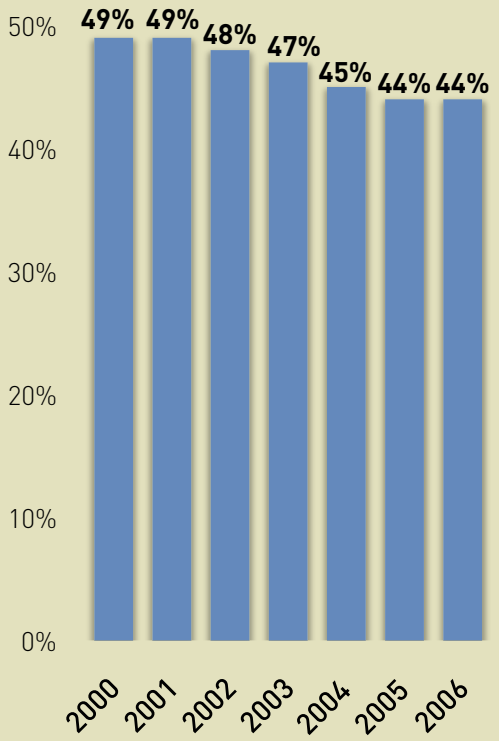
* Actual time may vary depending on working team composition, personnel dedicated to project, and complexity of analysis and implementation.

DEVELOPING BASELINE DISTRICT FACTS:
Cross-Sectional Analysis

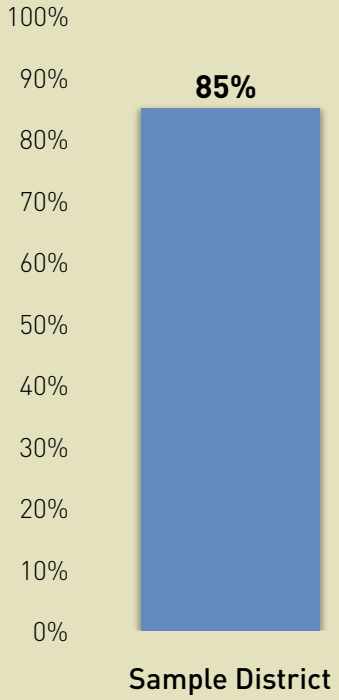
Key Definition

Cross-sectional analyses use data at a specific point in time or for a specific time period (e.g., the high school graduation rate in a given year).

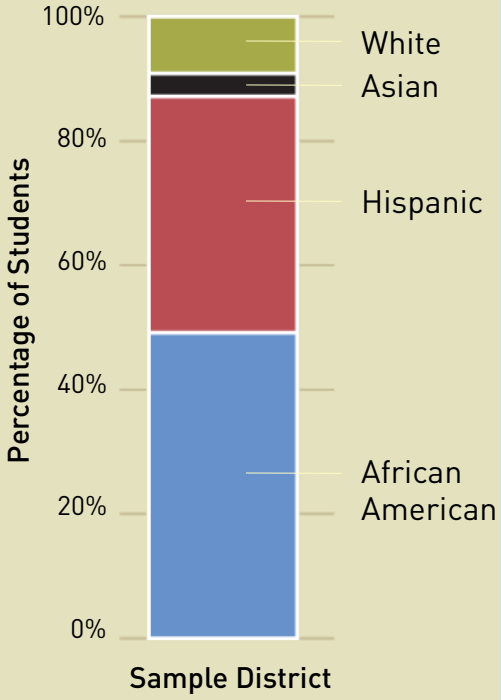
District High School Five-Year Dropout Rates, 2000–06



Percent Free and Reduced Lunch



Student Ethnicity



Key Steps

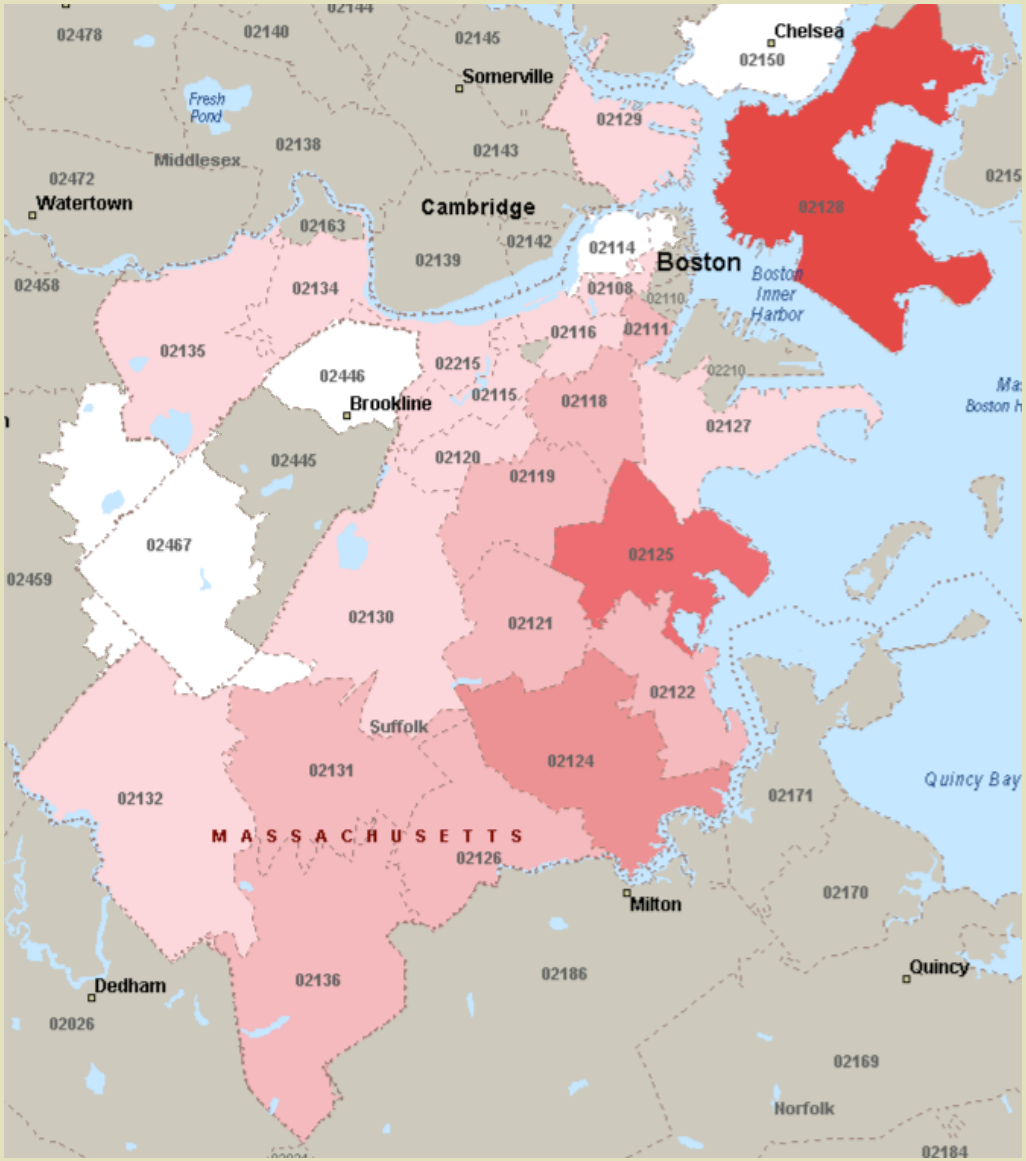
1. Create list of metrics that need to be agreed upon
2. Calculate these metrics based on the previously defined cohorts
3. Compare calculated metrics to those that already exist for the district (e.g., graduation rates) to ensure accuracy of methodology and cohort definition

Key Decisions/Questions

1. What facts are already known/ calculated by the district?
2. What common beliefs about the district need to be validated/refuted?

IDENTIFYING EFFECTIVE OPTIONS:
Geographical Concentrations of Student Need

Location of Overage Late Entrant ELL Students by Home Zip Code, 2005–06



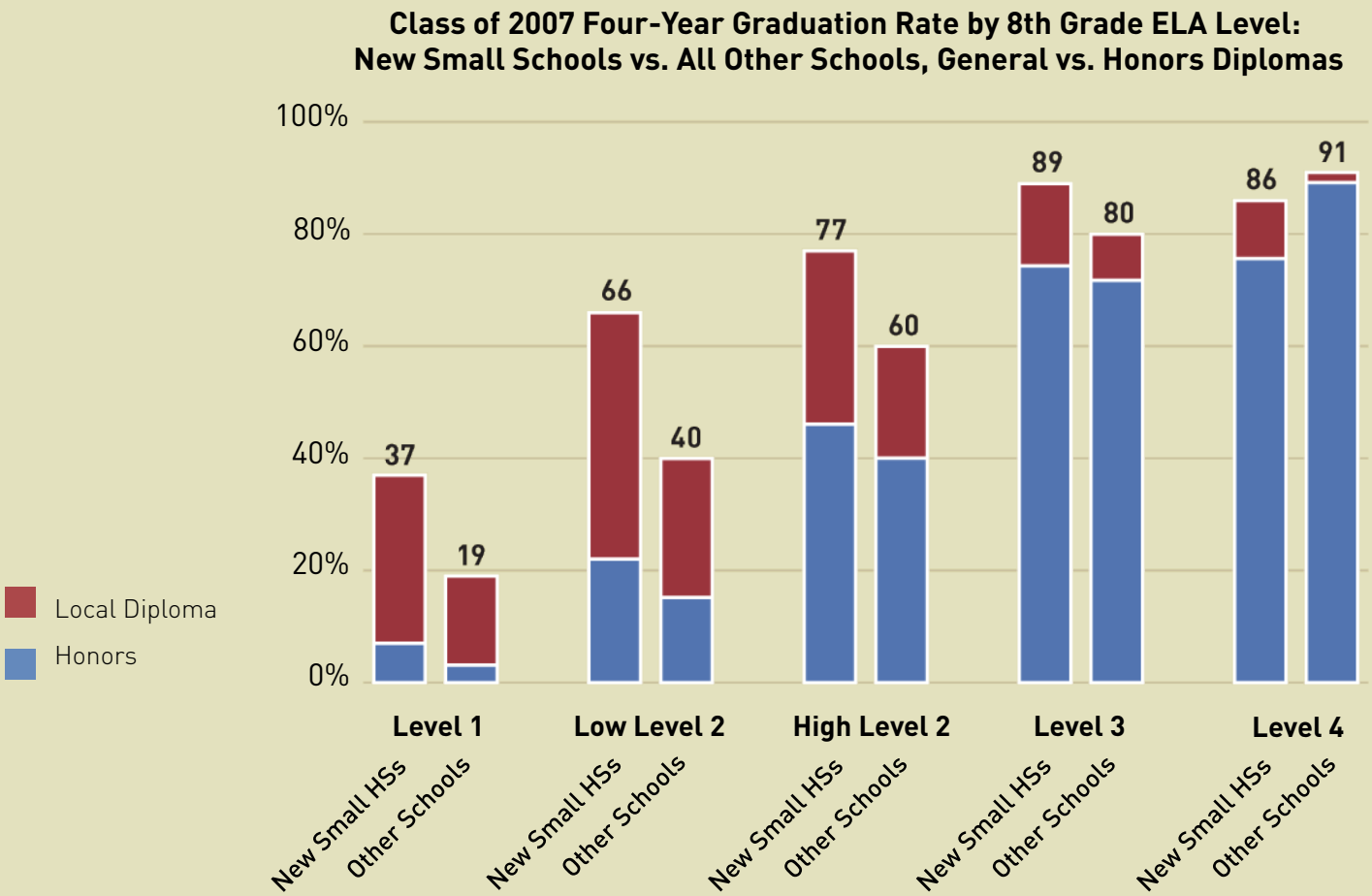
Key Steps

1. Decide what student need indicator is of interest
2. Use cross-sectional population definition similar to that in "sizing the off-track population"
3. Link students to home ZIP codes
4. Calculate percentage of students in total population that display indicator of need
5. Map high schools or programs focusing on this population against student need to assess whether need is being addressed

Key Decisions/Questions

1. What indicators of student need are of greatest interest?
2. What schools and programs currently exist that serve specific populations?

Determining Effect of Structure on Specific Students



	Level 1	Low Level 2	High Level 2	Level 3	Level 4
Overall Gap	18%	26%	17%	9%	N/A
Honors Gap	4%	7%	6%	3%	N/A

Key Learning

Controlling for specific student characteristics and comparing outcomes can provide a more nuanced view of performance without greatly increasing the difficulty of the analysis.

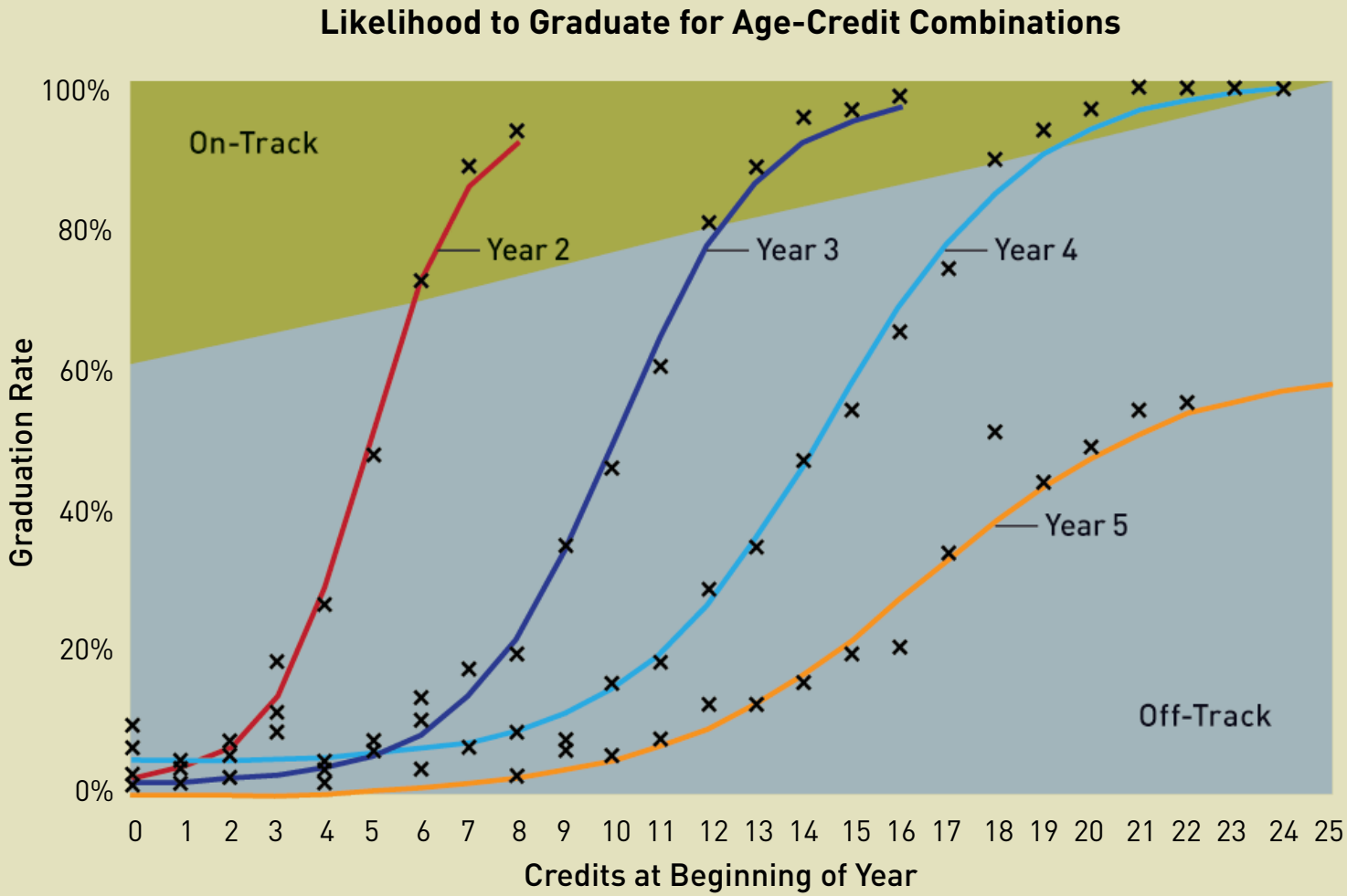
Key Steps

1. Use either cross-sectional or longitudinal student population, depending on questions to be answered
2. Determine what student characteristic will be controlled for (e.g., 8th grade test performance) and what school models are to be compared
3. Link students to schools based on previously established rule for when students link to schools
4. Calculate outcomes for each distinct student-school grouping

Key Decisions/Questions

1. Does analyzing school outcomes in this way sufficiently account for student-level differences?

Developing Basic Predictive Models



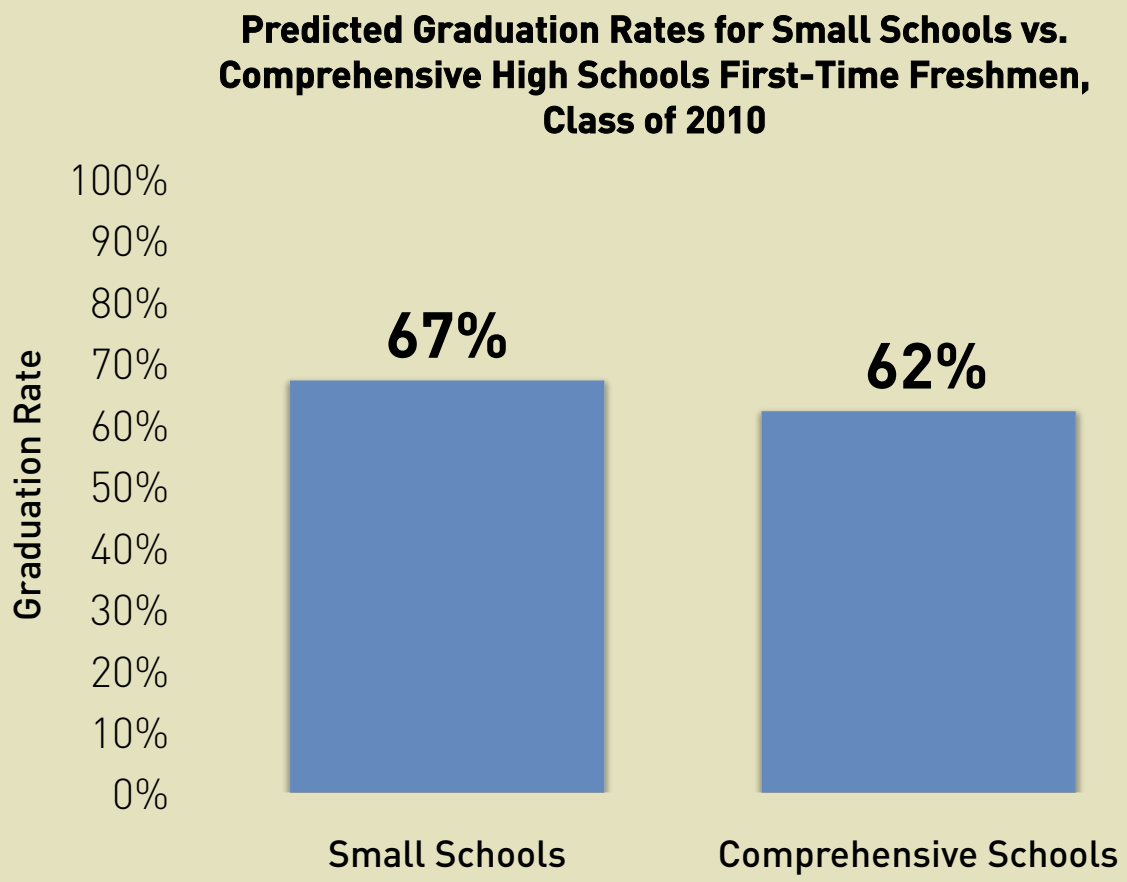
Key Steps

1. Use a longitudinal student population representing a single cohort of students
2. For each student, determine the number of credits earned each year and the enrollment status at the end of the year (e.g., still enrolled, graduate, drop out, transfer)
3. Calculate the graduation rate for each credits-earned group for each year in time

Key Decisions/Questions

1. When should total credits earned be calculated for a year? Including summer school credits?
2. Which graduation rate should be used—4 year, 5 year, 7 year?

Predicting Graduation Rates



Key Steps

1. Use historical relationship between credits earned and graduation rates (slide 6) to roughly project graduation rates of upcoming classes
2. For each student, calculate the number of credits. Based on student's credit accumulation and the historical graduation rate of students who have earned that credit level, assign student a graduation rate
3. Take a weighted average of each student to calculate total graduation rate

Key Decisions/Questions

1. What variables can be known early in a student's career that increase power of simple model?
2. What factors (i.e., promotion policy changes) may compromise the predictive power of the model?

Advanced Predictions of Performance: Regression Analysis

Key Learning

Not all regressions have significant school-level effects, as these effects may take several years to emerge.
Regressions using early outcomes like on-track may have negligible school-level factors.

Student-Level Factors

- Gender
- Ethnicity (White/Asian vs. African-American/Hispanic)
- Average 8th ELA and math proficiency level
- Age at entry to high school
- 8th grade attendance rate
- ELL status (Y/N)
- Special education (SPED) status (LRE vs. self-contained vs. non-SPED)
- Lack of 8th grade attendance or 8th grade test scores

School-Level Factors

- School enrollment
- A series of concentration variables:
  - Percent of students with a Low Level 2 or below on either ELA or Math
  - Percent of students entering 9th grade overage
  - Percent of students with 8th grade attendance below 90 percent
- Tested but not significant: avg. teacher tenure, principal tenure, general education class size, SPED class size, facilities capacity utilization

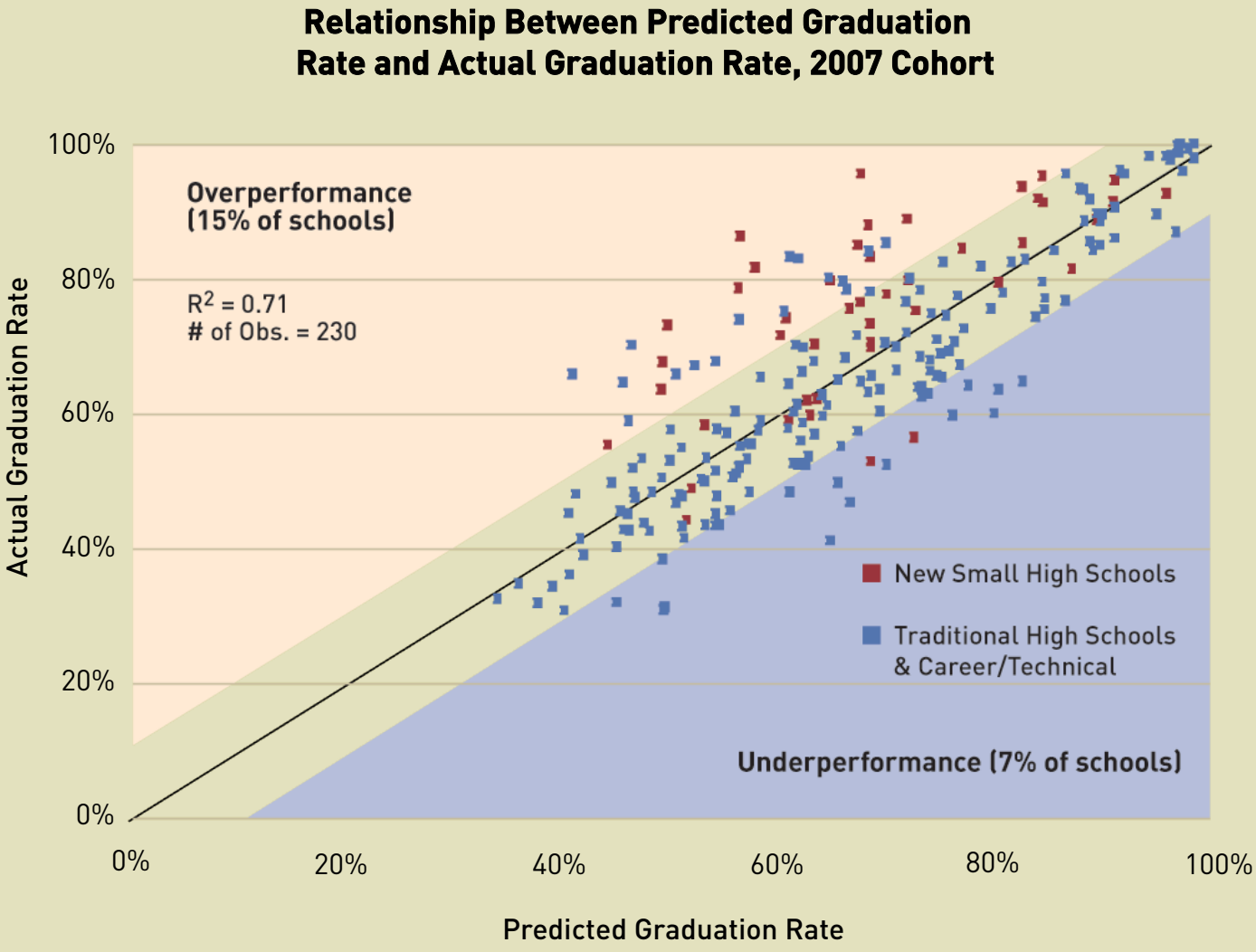
Key Steps

1. Determine outcome variable (e.g., on-track status, graduation)
2. Decide point in time to associate students to schools (suggested: freshman year school)
3. Assess variables that are known about students at entry (student-level) and are to be tested
4. Calculate size and concentration variables (school-level) that are to be tested

Key Decisions/Questions

1. Does the district have the internal expertise to develop a complex predictive model? If not, with whom can the district partner?

Using Multivariate Regression Analysis to Calculate Over- and Underperformance



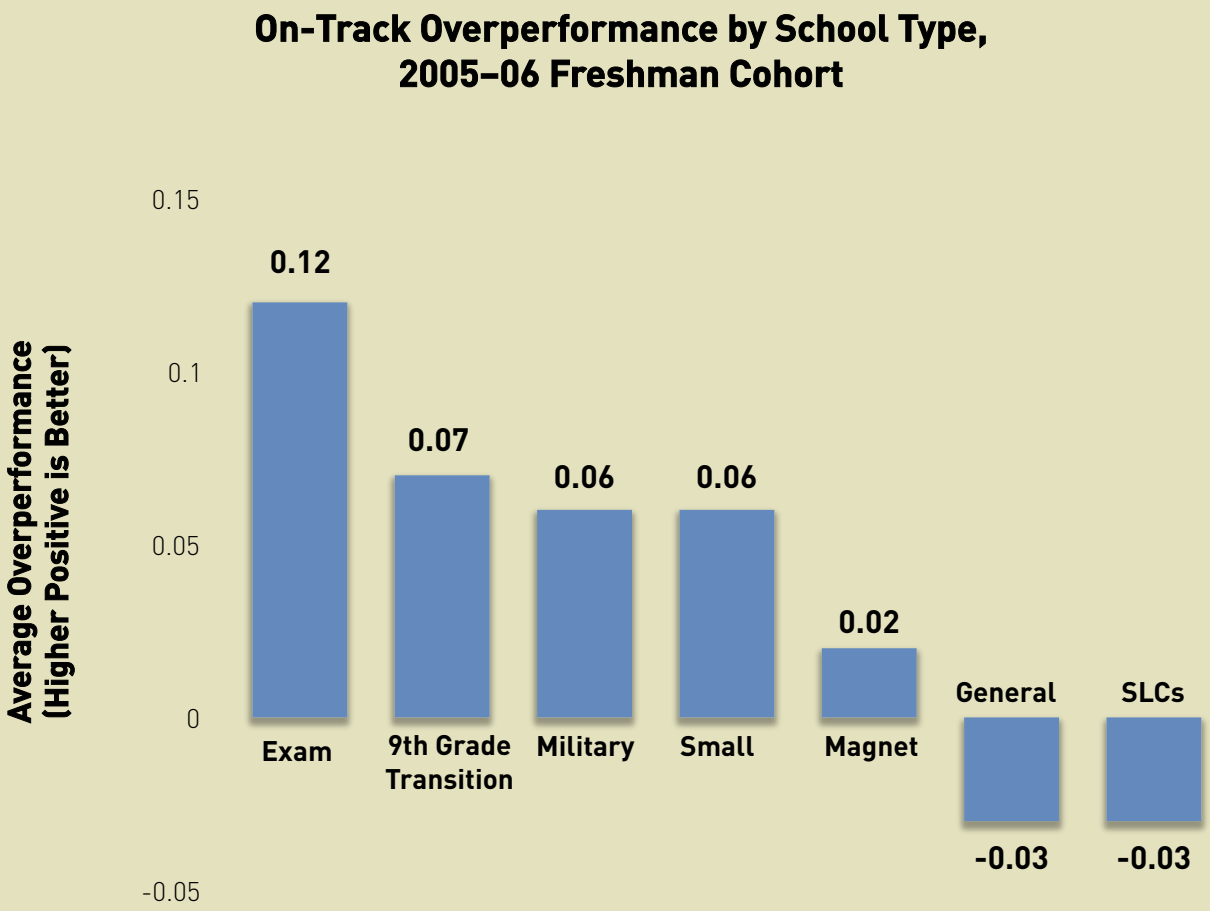
Key Steps

1. Use advanced predictive model methodology (slide 8)
2. Sum the individual student-level over/underperformance (actual – predicted) to determine school-level performance
3. Identify outcomes associated with different school models to determine whether school model produces systematic performance
4. Determine level of over/ underperformance that warrants further investigation

Key Decisions/Questions

1. Should over/underperformance be based on a single cohort's outcomes or the combination of multiple?

Identifying More Effective School Models



	Exam	9th Grade Transition	Military	Small	Magnet	General	SLCs
Enrollment	2.7K	1.0K	0.4K	2.5K	2.4K	14.9K	5.0K
% FRL	60%	92%	86%	92%	87%	89%	92%
% Overage	5%	80%	23%	34%	20%	29%	33%
% SPED	7%	4%	16%	21%	13%	17%	19%
% ELL	0%	0%	4%	3%	2%	6%	3%

Key Steps

1. Use advanced predictive model methodology (slide 8)
2. Group schools by model type
3. Average school-level residuals by model type to identify models that demonstrate consistently strong or weak performance
4. Developing profiles of school models may help provide context for the outcomes

Key Decisions/Questions

1. How do different school models support target student populations?

Understanding Model Variables under District Control

Variable	Description	What change would move the predicted graduation rate by 2 percentage points?
Total HS Enrollment	A high school's official grade 9–12 enrollment for the 2006–07 school year	A **500-student increase** (system avg=~1,500) in the total enrollment of a school will cause a decrease in graduation rate by 2 percentage points
Low 8th Grade Achievement	Percentace of students in a high school's 2007 cohort who scored Level 1 or Low Level 2 on at least one 8th grade exam (of students who took one or both 8th grade exams)	A **5 percentage point increase** (system avg=55%) in the number of "low 8th grade achievement" students in a school will cause a decrease in graduation rate by 2 percentage points
Overage at Entry	Percentage of students in a high school's 2007 cohort who turn 15 on or before 12/31/2003	A **5 percentage point increase** (system avg=31%) in the number of students in a school who were "overage at entry" will cause a decrease in graduation rate by 2 percentage points
Self-Contained (SC) Special Education (SPED) Proportion	Percentage of students in a high school's 2007 cohort who were ever SC SPED during grades 9-12	A **6 percentage point increase** (system avg=10%) in the number of SC SPED students in a school will cause a decrease in graduation rate by 2 percentage points

Key Steps

1. Separate schools from regression analysis into overperformers and underperformers
2. Select a performance level change to test (e.g., 2 percentage point change in graduation rate)
3. Identify variables that are substantially under district control—likely school size and student concentrations
4. Hold all other model variables constant while varying (solving for) a single variable to understand change

Key Decisions/Questions

1. What variables in the predictive model are substantially within district control of schools?
2. How does this analysis relate to schools that are "structured to fail"?

Understanding the Effect of Size and Concentration on Individual Students

School 1
- 400 students
- Concentrations of challenged students **15 percentage points below** system average

83% Graduation Rate

(Effect of 400 Student Enrollment Alone 78%)

"The Median Student"
- Female
- African-American/Hispanic
- Avg. 8th Grade Test Scores (2.7)
- Avg. 8th Grade Attendance (92%)
- On-Age at Entry to HS (14 yo)
- Neither ELL nor SPED

70% Predicted Graduation Rate

School 2
- 800 students
- Concentrations of challenged students **5 percentage points below** system average

78% Graduation Rate

School 3
- 1,500 students
- Concentrations of challenged students **5 percentage points above** system average

70% Graduation Rate

School 4
- 3,000 students
- Concentrations of challenged students **15 percentage points below** system average

55% Graduation Rate

Key Steps

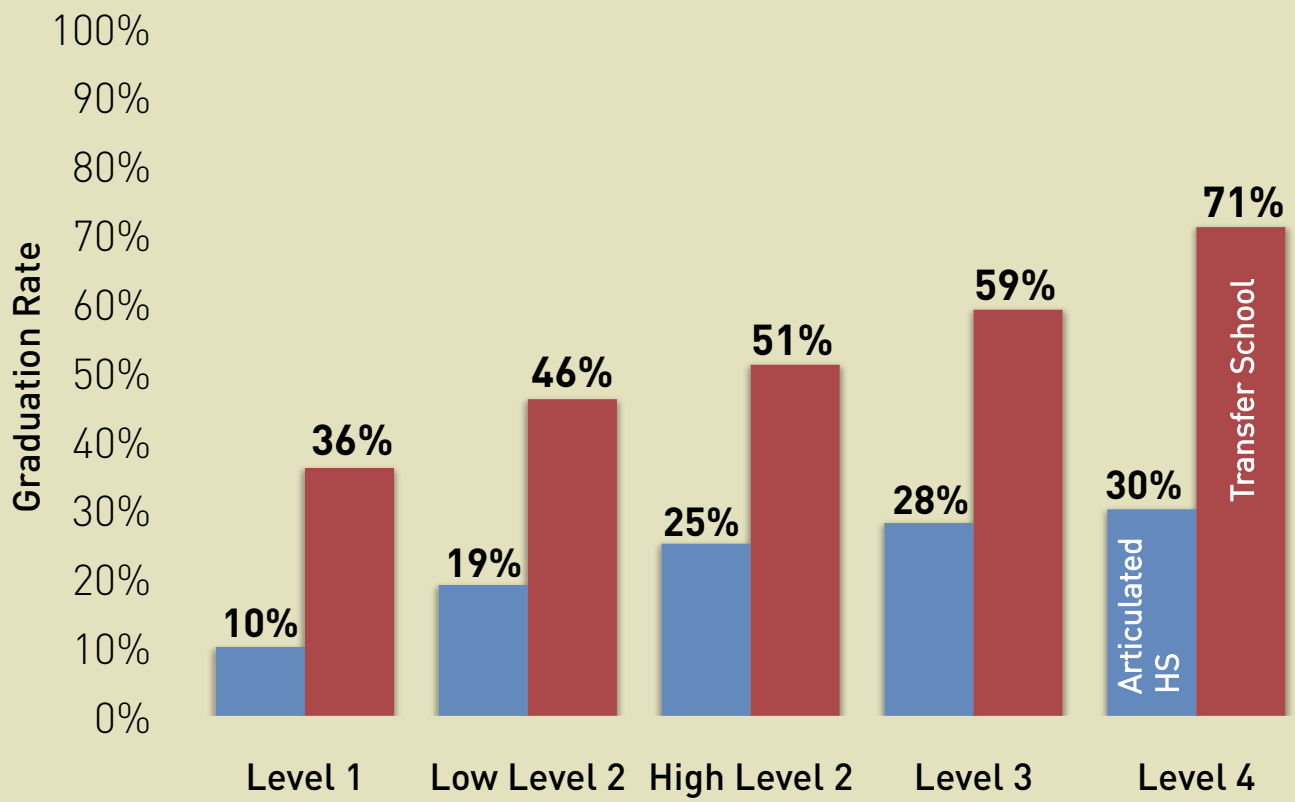
1. Select a representative student—in the example, the student has the median characteristics of all students—and determine the graduation rate independent of school structure
2. Choose several school structures that are representative of the variety of schools in the district, with different size and concentrations
3. Adjust model size and concentration parameters to calculate student graduation rate given different school structures

Key Decisions/Questions

1. For representative student, how much of a performance gain is possible with just structural changes?

Determining Differential Effectiveness

Overage and Under-Credited Seven-Year Graduation Rates by Reading Level: Articulated High Schools vs. Transfer Schools



	Level 1	Low Level 2	High Level 2	Level 3	Level 4
Improvement vs. General HS	278%	145%	109%	111%	137%

Key Steps

1. Use either cross-sectional or longitudinal student population, depending on questions to be answered
2. Determine what student characteristic will be controlled for (e.g., 8th grade test performance) and what school models are to be compared
3. Link students to schools based on previously established rule for when students link to schools
4. Calculate outcomes for each distinct student-school grouping

Key Decisions/Questions

1. How does student performance across key metrics vary by school type?

BILL & MELINDA GATES foundation

Information for this report is drawn from Atlanta, Boston, Charlotte-Mecklenburg, Chicago, Dallas, Portland, and New York City and their partnerships with the Bridgespan Group, Boston Consulting Group, McKinsey & Company, Education Resource Strategies, and The Parthenon Group.