

# Targeting Education Interventions For Young Students

Eric & Wendy Schmidt  
Data Science For Social Good  
Summer Fellowship



**Max Klein**  
GroupLens Research  
University of Minnesota

**Monica Alexander**  
University of California,  
Berkeley

**Charlotte Huang**  
University of Chicago

**Chad Kenney**  
University of Chicago

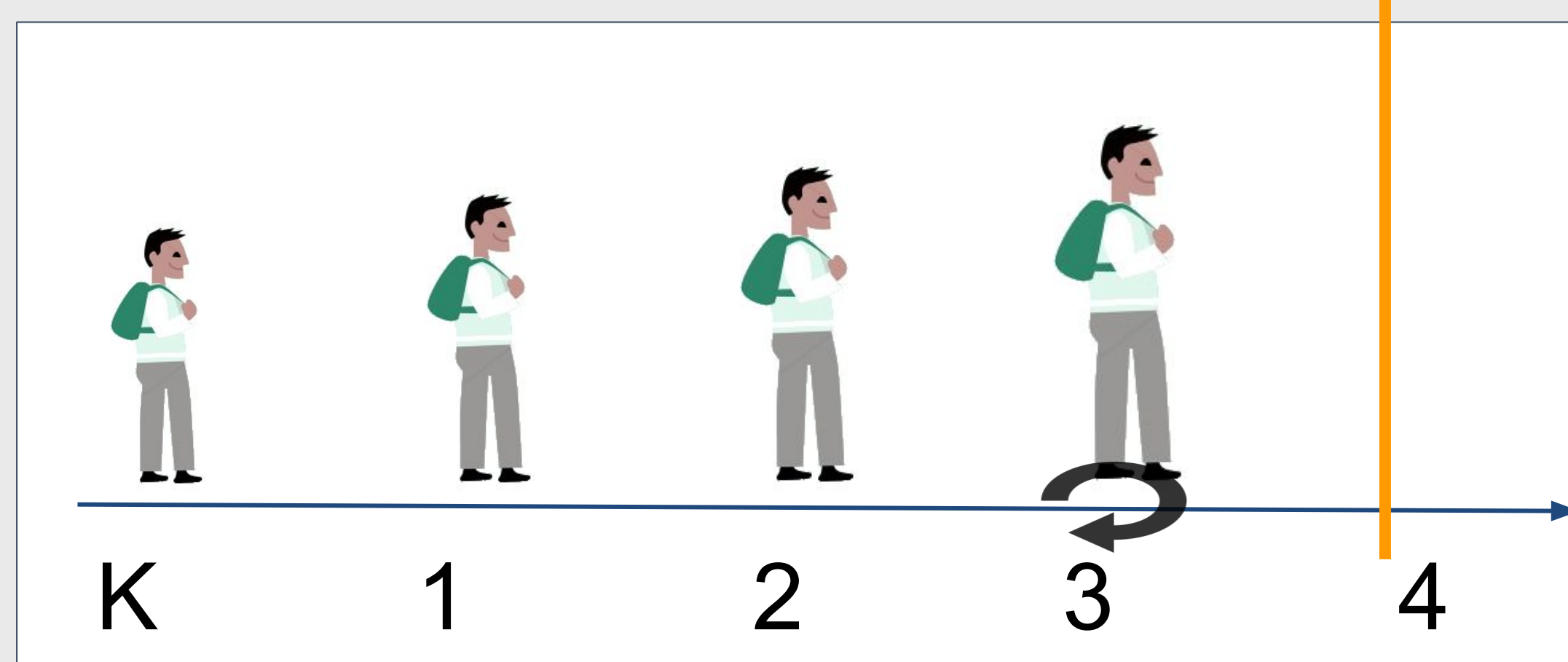
**R. Ali Vanderveld**  
University of Chicago

**Kevin H. Wilson**  
University of Chicago

1

## Introduction

Third grade is often thought to be a critical time where students transition from **learning to read** to **reading to learn**. In Oklahoma, the **Reading Sufficiency Act (RSA)** of 2013 requires less proficient students to repeat third grade. This creates a **high-stakes testing environment**.

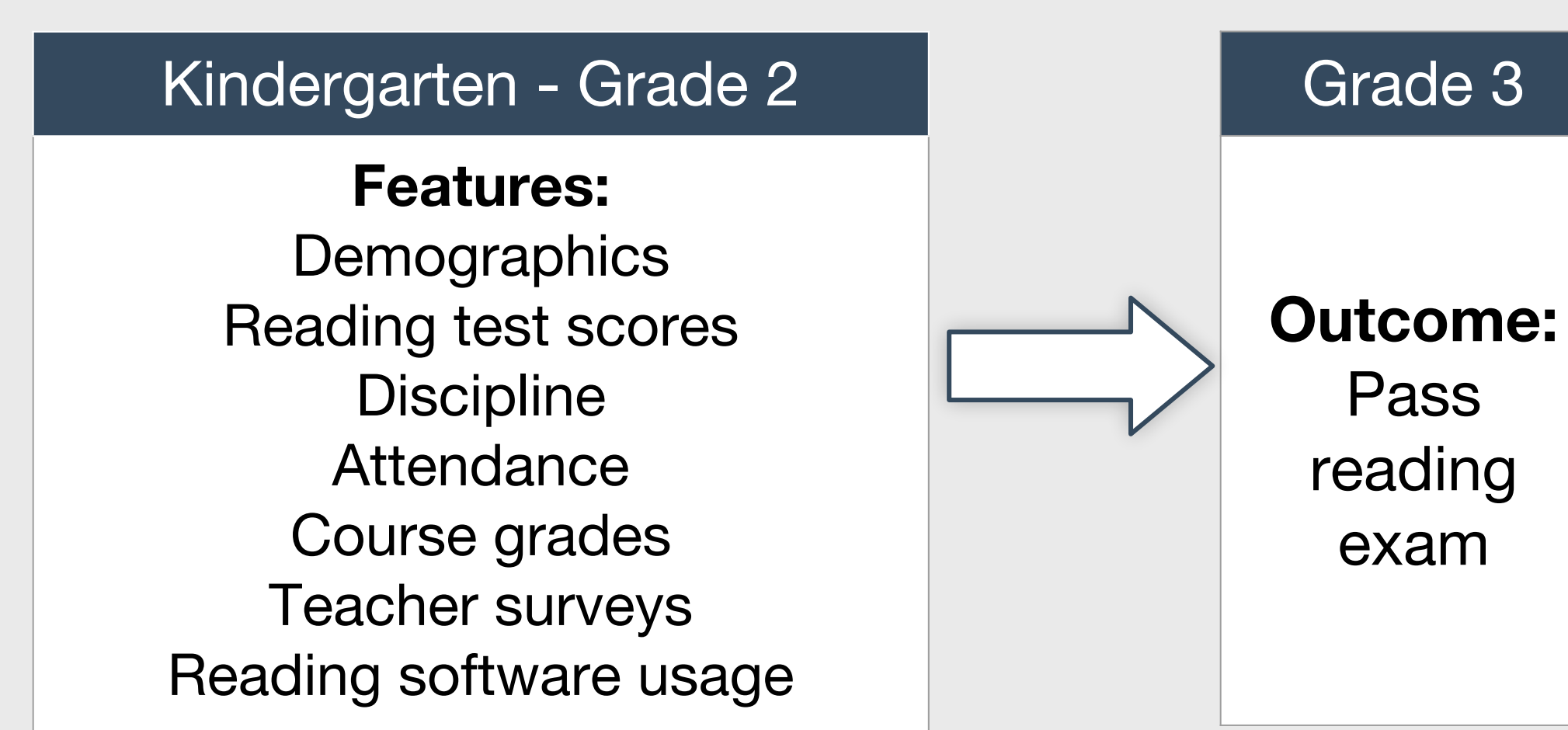


3

## Methods

We applied machine learning models to deliver data-driven predictions of third grade reading proficiency that are accurate and interpretable.

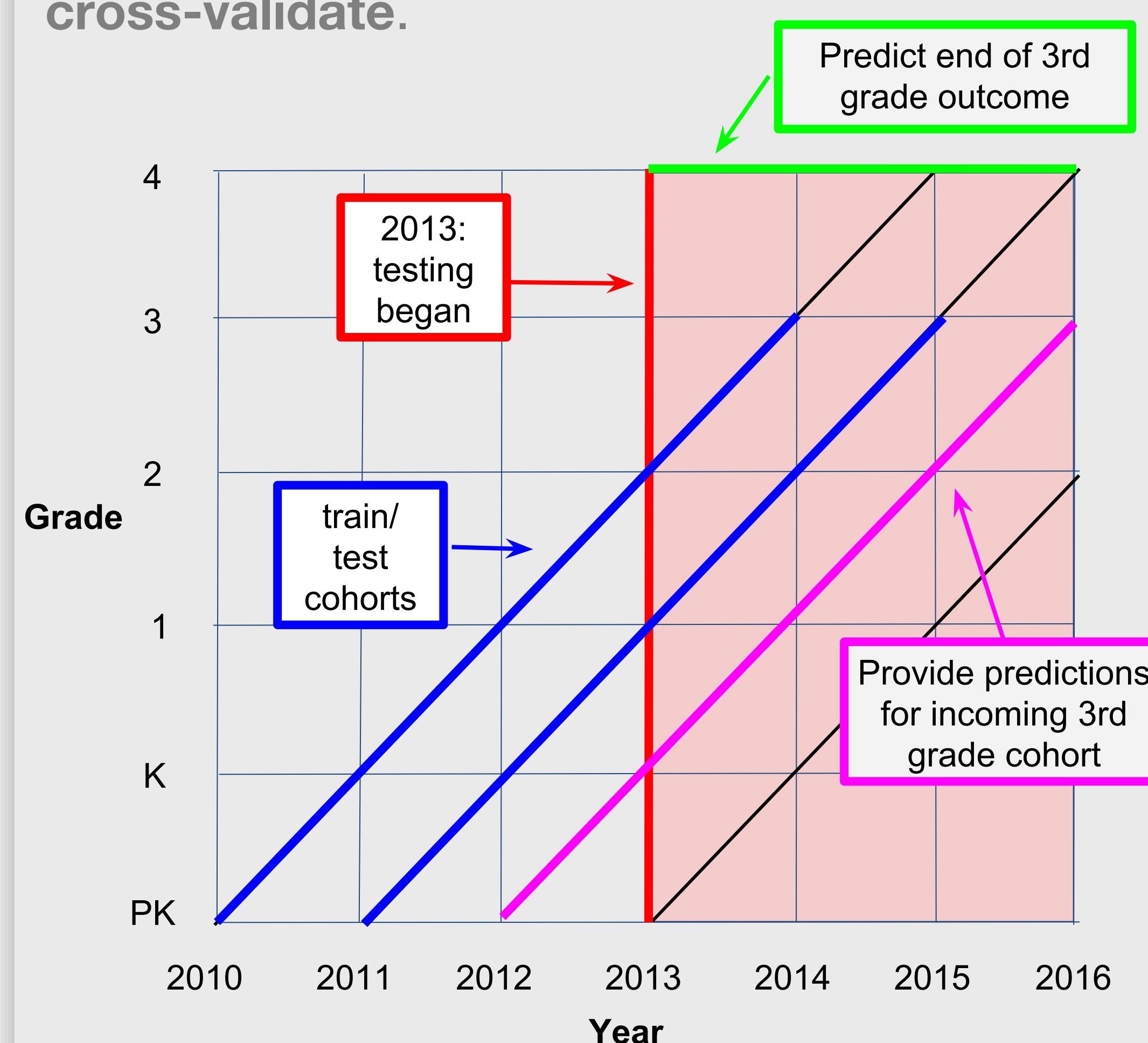
### Model



Based on a longitudinal student dataset, we created an **end-of-second-grade model** to determine **whether a student will pass a third-grade-level reading exam** in the next year.

### Model Validation

Since the passage of RSA in 2013, there have been two complete cohorts which have taken reading tests. We **split** the data into training and tests sets **between these two cohorts to cross-validate**.



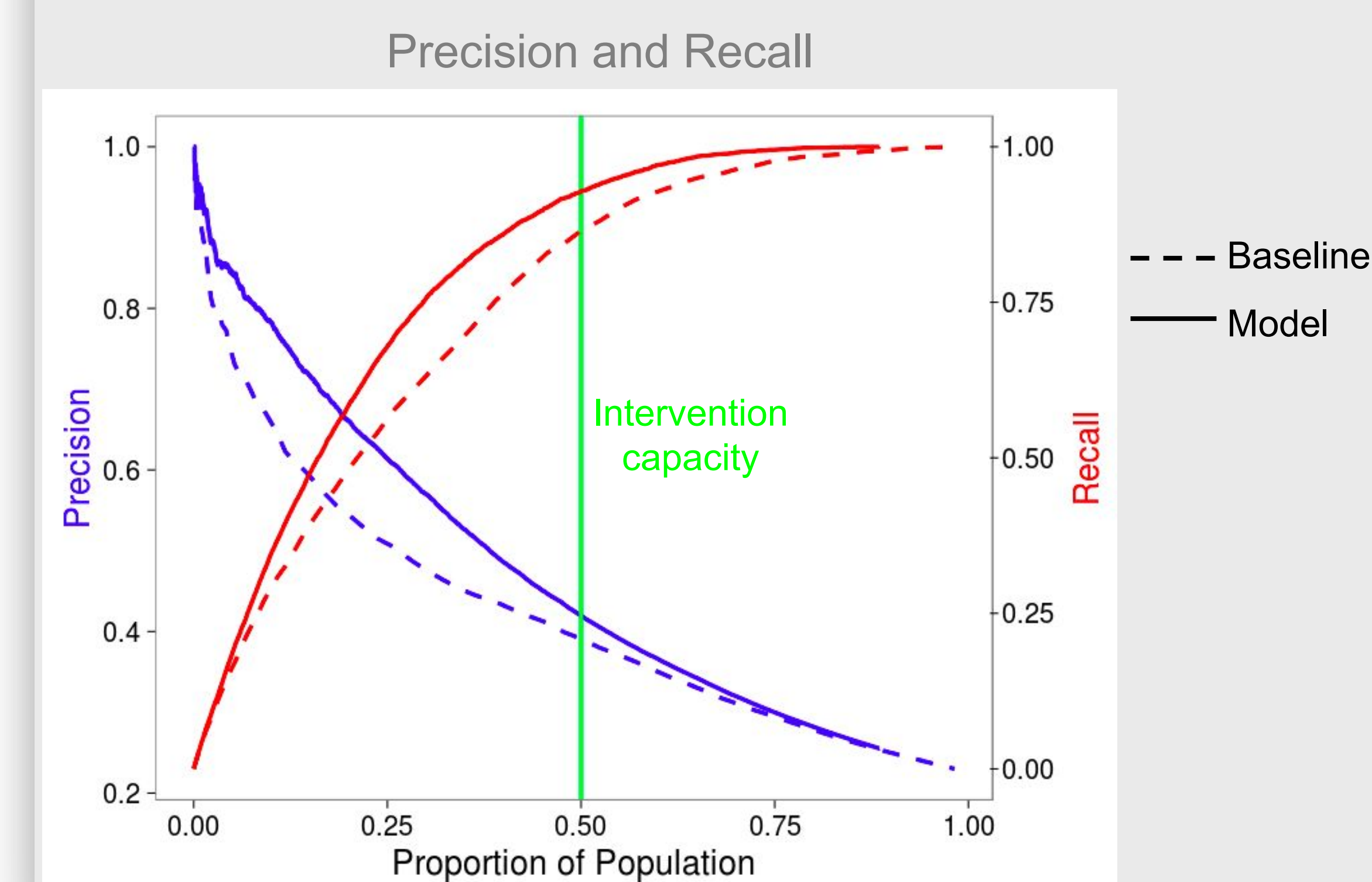
### Classifier Performance

After testing a wide range of classification methods, we found that Random Forest and penalized Logistic Regression give the best performance in terms of prediction precision and recall.

4

## Results

Tulsa Public Schools has the capacity to **provide extra reading help to approximately half of students** in any particular year. At this level, results from our best performing model indicate an ability to correctly **recall 95%** of students who are not on track. This is an improvement on the baseline method (using past test score results alone) by around 5 percentage points.



5

## Impact

Our increase in recall implies **250 more struggling students** every year will be given extra attention. This intelligence will be provided to administrators and schools through a **risk dashboard for teachers**. We hope this will prevent many young students from needing to repeat third grade—and help improve their life trends overall.

