UNIVERSITY OF CALIFORNIA

Los Angeles

Causal Effects of Poverty Based Funding Thresholds on Educational Outcomes: A Gaussian Process Regression Discontinuity Analysis of Los Angeles Unified School District

A thesis submitted in partial satisfaction

of the requirements for the degree

Master of Applied Statistics and Data Science

by

Devin Nathaniel Reeh

2025

ABSTRACT OF THE THESIS

Causal Effects of Poverty Based Funding Thresholds on Educational Outcomes: A Gaussian Process Regression Discontinuity Analysis of Los Angeles Unified School District

by

Devin Nathaniel Reeh
Master of Applied Statistics and Data Science
University of California, Los Angeles, 2025
Professor Chad Hazlett, Chair

This thesis examines the causal effects of poverty based eligibility thresholds on school level outcomes in the Los Angeles Unified School District (LAUSD). The analysis focuses on two administrative metrics used to allocate resources and services: Free and Reduced-Price Meals (FRPM) eligibility and Unduplicated Pupil Percentage (UPP). Leveraging 35 percent 55 percent, and 75 percent levels as policy based cutoffs, the study applies both classical Regression Discontinuity Design (RDD) and its Bayesian extension, Gaussian Process Regression Discontinuity Design (GP-RDD), to estimate local treatment effects.

Using administrative data from the 2021-2022 academic year, this research evaluates the effects of qualifying for supplemental funding support on chronic absenteeism, standardized test performance in English Language Arts and Mathematics (ELA), and participation in the Beyond the Bell (BTB) after-school program. Effects on chronic absenteeism are small and statistically insignificant at both the 35% and 75% FRPM thresholds. Under classical RDD at 35%, the point estimate is +6.48 percentage points (95% CI: $[-1.78, 14.74]$, $p = 0.124$),

and under GP–RDD it is +7.87 percentage points (95% CrI: $[-3.38, 19.11]$, $p = 0.171$). At 75%, estimates are near zero and non–significant across methods.

The thesis of Devin Nathaniel Reeh is approved.

Rick Schoenberg

Michael Tsiang

Chad Hazlett, Committee Chair

University of California, Los Angeles

2025

To Kelsea Valerio, my partner, whose unwavering emotional support carried me from start to finish.

And to my parents, Daphne and Edmund Reeh, who made this possible through their sacrifices, steadfast belief, and financial support.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

CHAPTER 1

Introduction

1.1 Research Problem

Public education systems in the United States continue to struggle with closing achievement gaps for students from economically disadvantaged backgrounds. In the Los Angeles Unified School District (LAUSD), one of the nation's largest and most demographically diverse school systems, resources and support are often distributed based on school level poverty indicators such as Free and Reduced-Price Meal (FRPM) eligibility and Unduplicated Pupil Percentage (UPP), the latter used under California's Local Control Funding Formula (LCFF).

Schools that cross administrative thresholds on these indicators may become eligible for additional funding, such as Title I allocations at 35% or 75% FRPM, or for state level grants triggered at 55% or 75% UPP. Programs like *Beyond the Bell* (BTB), LAUSD's after school initiative, are frequently prioritized for high poverty schools, although access is contingent on grant funding, staffing, and district discretion. Since the use of these cutoffs to allocate opportunity and support is widespread, this thesis seeks to estimate the causal effects of these programs on student outcomes by using these cutoffs.

While existing research has evaluated some threshold based policies in different geographies, most studies focus on broad national programs or rely on parametric methods that impose rigid modeling assumptions. Few examine LAUSD specifically, and even fewer employ flexible, nonparametric designs.

This thesis seeks to estimate the effects of key support programs that are allocated using poverty thresholds in LAUSD. We use the administrative cutoffs (e.g., FRPM 35%, 75%; UPP 55%, 75%) as an identification strategy to causally identify local program impacts while minimizing confounding near the boundaries. To accomplish this, we use polynomial Regression Discontinuity Design (RDD) and its nonparametric Bayesian extension, Gaussian Process RDD (GP-RDD). By exploiting quasi-random assignment near commonly used policy cutoffs, this approach identifies the localized impacts of program eligibility in one of the most policy-relevant districts in California.

1.2 Background and Motivation

In California's K–12 funding system, poverty indicators such as FRPM eligibility and UPP are central to determining how resources are distributed. These cutoffs drive federal Title I allocations, state LCFF supplemental and concentration grants, and eligibility for initiatives such as the Community Eligibility Provision (CEP), which expands school meal access. In addition, LAUSD programs such as BTB, an extended before/after school program, often use these same thresholds as funding criteria which heavily influence whether schools will offer the after school programs.

Eligibility thresholds determine whether a school receives additional support, which in turn may influence outcomes. For example, schools with UPP above 75% are given higher priority for larger concentration grants under LCFF, and schools above 35% FRPM are typically eligible for Title I funded interventions.

This institutional setup allows the use of regression discontinuity designs to study the causal impact of programs and funding, leveraging the thresholds to obtain unbiased estimates. This study focuses on outcomes that are sensitive to resource constraints, including attendance, standardized test performance, and access to extended learning time. It investigates whether policy thresholds function as meaningful levers for improving student outcomes.

1.3 Scope

This thesis focuses on school level outcomes in LAUSD for the 2021-2022 academic year. It uses the FRPM 35%, FRPM 75%, UPP 55%, and UPP 75% in order to estimate the causal effects of programs and funding accessed through eligibility thresholds on three types of outcomes: (1) chronic absenteeism, (2) participation in the Beyond the Bell after-school program, and (3) performance on CAASPP standardized tests in English Language Arts and Math.

Methodologically, the analysis uses sharp RDD and Gaussian Process RDD to estimate local treatment effects near policy cutoffs. These designs focus on identification at the margin of program eligibility and do not attempt to estimate average effects across the entire poverty distribution. The study is observational and cross-sectional, relying on administrative data from the 2021-2022 school year, with some lagged outcomes from prior years used as covariates.

1.4 Key Contribution

This thesis makes four primary contributions:

1. It applies a flexible, nonparametric Bayesian method (GP-RDD) to evaluate school level policy thresholds in a real world education setting.

2. It compares results across modeling strategies, including local linear and polynomial RDDs, to assess how effect estimates vary with functional form assumptions.

3. It estimates localized impacts of FRPM and UPP thresholds on behavioral and academic outcomes, providing new evidence on the effects of qualifying for support on student outcomes.

4. It contributes a district level case study from LAUSD, offering insights relevant to both

California education policy.

CHAPTER 2

Literature Review

2.1 Introduction

This literature review summarizes empirical research across three domains relevant to the thesis: (1) school nutrition programs, (2) after-school programs, and (3) causal inference methods used to evaluate education policies. These programs represent two of the most widely implemented support strategies in K–12 education, particularly in high poverty urban districts like LAUSD. While substantial research exists on each intervention individually, less is known about how eligibility criteria based on administrative poverty thresholds, such as Free and Reduced-Price Meal (FRPM) eligibility and Unduplicated Pupil Percentage (UPP), shape school level outcomes in LAUSD. This review highlights established findings and identifies a key gap in the literature that this thesis seeks to addresss.

2.2 School Nutrition Programs and Student Outcomes

A growing body of research has assessed the effects of school nutrition programs, particularly universal free lunch and breakfast, on academic and behavioral outcomes. Schwartz and Rothbart (2024) use a difference-in-differences (DiD) approach to examine New York City's expansion of universal free lunch to all middle schoolers and find modest improvements in test scores. Gains were concentrated among non-poor students, with estimated increases of 0.08 standard deviations in math and 0.06 in English Language Arts (ELA) [SR20].

Leos-Urbel et al. (2013) also use DiD to evaluate the rollout of universal free breakfast in New York City. They report limited increases in participation and no significant effects on attendance or academic outcomes [LSW13]. These findings suggest that delivery logistics and school level implementation may affect the program's impact.

Other research explores alternative models such as "Breakfast After the Bell" (BAB), in which meals are served during instructional time. Cuadros-Meñaca et al. (2021) find a modest negative effect on math achievement in Arkansas and no effect in ELA [CTN22]. In contrast, Bleeker et al. (2012), studying a large urban district, find achievement gains of approximately 0.10 standard deviations in math and reading, particularly for low income and Hispanic students [Ble12]. These contrasting results highlight the importance of accounting for implementation heterogeneity and local context.

Ritchie et al. (2024) synthesize six studies on universal free school meal (UFSM) programs and find consistent increases in meal participation, but only small or null effects on attendance and academic performance [Rit24]. Currie and Moretti (2003) provide a theoretical motivation for these investments by linking early-life public programs to long-term human capital formation [CM03]. Gundersen and Ziliak (2014) similarly emphasize the role of food security in supporting educational progress [GZ14].

While this literature provides useful insight, most studies focus on national or East Coast contexts and use DiD designs. Few apply RDD or analyze post-2018 expansions in California, and none directly examine the consequences of narrowly meeting or missing poverty based thresholds like those used in California's Local Control Funding Formula. These gaps motivate the use of localized threshold based designs in this thesis.

2.3 after-school Programs and Student Outcomes

Well designed after-school programs have shown potential to improve student outcomes, but results vary depending on implementation and study design. Jenson et al. (2023)

evaluate a community based after-school program serving low income neighborhoods using a matched comparison design. They find that participation improves reading scores, increases attendance, and reduces suspensions. However, the study is limited by potential selection bias [Jen23].

Vandell et al. (2007) study LAUSD's LA's BEST program using matched cohorts and find modest gains in test scores and substantial reductions in juvenile offenses among consistent participants [VRP07]. These results underscore the importance of sustained participation and program fidelity. However, the study does not use quasi-experimental methods, limiting causal interpretation.

Dynarski et al. (2004), in a federal evaluation of the 21st Century Community Learning Centers program, use DiD to assess after-school programming in multiple states. They find limited academic gains but report improvements in behavior and student safety [Dyn04]. Durlak et al. (2010), through meta-analysis, show that programs with structured curricula and explicit skill building components yield more consistent improvements in academic and behavioral metrics [DWP10].

Although there is growing evidence on after-school programming, most studies evaluate program impact rather than the consequences of eligibility for access. In districts like LAUSD, participation in programs such as *Beyond the Bell* is often prioritized for schools exceeding specific FRPM or UPP thresholds. However, there is limited evidence on whether these programs effect actual participation rates or downstream outcomes. This thesis addresses that gap by estimating the local effects of eligibility on school level after-school program access and associated outcomes.

2.4 Causal Inference Methods for School based Interventions

Quasi-experimental methods have become essential tools for evaluating education policies in real world settings where randomized controlled trials are impractical. Core approaches include DiD, instrumental variables, and panel fixed effects, each of which attempts to control for confounding under specific assumptions. However, these methods can be sensitive to violations such as nonparallel trends or endogenous policy adoption.

Recent advancements address these challenges. Callaway and Sant'Anna (2021) offer a generalized DiD estimator that accommodates treatment heterogeneity and staggered rollout [CS21]. Goodman-Bacon (2021) shows how traditional two way fixed effects can produce biased estimates when treatment timing varies across units [Goo21]. Athey and Imbens (2022) emphasize the importance of pre-analysis plans and design based identification in applied work [AI22].

RDD has emerged as a particularly relevant tool for studying education policies that assign treatment based on fixed thresholds. Van der Klaauw (2008) uses RDD to assess Title I funding in New York City and finds that although the program increased resources, it did not improve academic outcomes [Kla08]. Jacob and Lefgren (2004) use test score cutoffs to evaluate summer school and grade retention in Chicago, reporting positive effects for younger students [JL04]. Angrist and Lavy (1999) study class size reductions in Israel using an enrollment based threshold and find significant gains in test performance [AL99].

Other RDD applications include Sohn et al. (2022), who study a performance based funding policy in South Korea and find improved student achievement [Soh22]. Schwerdt, West, and Winters (2017) analyze Florida's third-grade retention policy and observe short-term test score gains that later diminish [SWW17]. Manacorda (2012) finds that grade retention in Uruguay increases dropout risk [Man12]. Canbolat (2023) applies multi-cutoff RDD to an attendance based early warning system, finding reductions in chronic absenteeism for higher SES students [Can23].

While these studies demonstrate the utility of RDD for estimating credible local treatment effects, few apply this framework to administrative poverty thresholds like FRPM and UPP. Fewer still use these thresholds to study eligibility for multiple programs simultaneously. This thesis contributes to the literature by using school level FRPM and UPP thresholds in LAUSD to estimate the causal effects of program funding eligibility on student behavioral and standardized academic outcomes. It also employs Gaussian Process RDD (GP-RDD), a Bayesian extension of RDD that flexibly models treatment effects when data are sparse near the cutoff. GP-RDD avoids rigid polynomial assumptions and allows for uncertainty quantification, making it well-suited for education policy applications involving small samples or complex functional forms.

CHAPTER 3

Institutional Context: Funding and Program Access in LAUSD

3.1 Overview of LAUSD

The Los Angeles Unified School District (LAUSD) is the second largest public school district in the United States, serving more than 430,000 students across over 1,000 schools within a 710 square mile area that includes the city of Los Angeles and 31 surrounding municipalities. As of the 2023–2024 school year, LAUSD's student population is approximately 73% Latino, 10% White, 7% Black, and 4% Asian, with more than 86% classified as socioeconomically disadvantaged and 23% identified as English learners. This scale and diversity make LAUSD a nationally significant site for studying equity driven education policies.

3.2 The Local Control Funding Formula (LCFF)

California funds its public K–12 schools through the Local Control Funding Formula (LCFF), implemented in 2013. LCFF allocates funds based on student need, using the following components:

- **Base grants**: Allocated for all students based on grade level.
- **Supplemental grants**: Provided for each *unduplicated* student who is low income, an English learner, or a foster youth.

- **Concentration grants**: Awarded to districts and schools with more than 55% unduplicated students, with intensified funding above 75% [Edu21a].

3.2.1 Unduplicated Pupil Percentage (UPP)

The Unduplicated Pupil Percentage (UPP) is calculated annually and includes students qualifying for at least one high need category. The two most policy relevant thresholds are:

- **55% UPP**: Triggers supplemental funding for each unduplicated student [Edu21a].
- **75% UPP**: Prioritizes schools for additional LCFF funds and expanded learning supports, including the Expanded Learning Opportunities Program (ELO-P) [Inf21].

These thresholds represent administrative discontinuities in funding intensity and are used in causal estimation strategies throughout this thesis.

3.3 FRPM Percentage as a Poverty Proxy

Free and Reduced Price Meal (FRPM) eligibility is widely used as a proxy for school level poverty. Eligibility is based on federal income criteria (185% of the poverty line) and reported annually by each school. Though distinct from UPP, FRPM% is strongly correlated and often used to assign funding and program access [Edu21b].

3.3.1 Key FRPM Thresholds

FRPM thresholds serve as eligibility gates for major federal and state programs:

- **35%**: Title I eligibility and participation in the Community Eligibility Provision (CEP) [Edu21b].
- **55%**: Aligns with the LCFF supplemental grant threshold [Edu21a].

- **75%**: Priority threshold for after-school programs and full CEP reimbursement [Inf21].

3.4 Program Access Driven by Thresholds

The following programs rely on FRPM or UPP cutoffs to determine eligibility or prioritization:

- **Title I**: Academic support funds allocated beginning at 35% FRPM [Edu21b].
- **CEP (Community Eligibility Provision)**: Universal meal access begins at 35% identified student percentage [Edu21b].
- **ASES and 21st CCLC**: after-school programs prioritized for schools above 75% FRPM or UPP [Inf21].
- **ELO-P (Expanded Learning Opportunities Program)**: Priority given to schools with UPP above 75% [Inf21].

Schools that cross these thresholds receive access to additional instructional time, staffing, and operational support.

3.5 Discontinuities as Opportunities for Causal Inference

FRPM and UPP thresholds determine access to specific programs and funding streams. By comparing schools just above and below these thresholds, we can estimate the causal effects of the programs and funding provided, while minimizing confounding..

3.6 Summary of Policy Relevant Thresholds

Threshold	**Used For**	**Source**
35% FRPM	Title I eligibility, CEP	US DOE Title I Guidance, USDA CEP Manual, [Edu21b]
75% FRPM	Federal Title I eligibility	US DOE Title I Guidance, USDA CEP Manual, [Edu21b]
55% UPP	LCFF supplemental funding	CDE LCFF Overview, [Edu21a]
75% UPP	Prioritization for ASES, ELO-P	CDE Allo Page, California Ed Code, [Inf21]

Table 3.1: Policy relevant administrative thresholds and their associated programs.

CHAPTER 4

Data and Variable Construction

4.1 Data Sources

The dataset is a school level panel covering academic year 2021-2022 for all public K-12 schools in LAUSD. Data were merged from publicly available sources maintained by the California Department of Education (CDE) and LAUSD. Each data source and its provider are listed below.

Data Source & Hyperlink	**Description**
CALPADS Unduplicated Pupil Count File	Unduplicated pupil counts and subgroup breakdowns
Chronic Absenteeism File	school level absenteeism rates
Enrollment by School	Enrollment by school and grade, 2020–2022
FRPM Data	Annual FRPM eligibility rates
BTB Program Participation	Before/after school participation in LAUSD
CAASPP Results	ELA/Math proficiency rates by school

Table 4.1: Data sources used to construct the dataset

4.2 Data Linkage and Cleaning

Multiple datasets were joined using unique school codes from the California Department of Education and LAUSD. Datasets were cleaned and merged in R using a modular pipeline.

Handling Missing Values and Suppressed Cells

CDE suppresses records where fewer than 11 students are represented, denoting them with '*'. These were converted to `NA`. When possible, chronic absenteeism percentages were re-calculated:

$$\text{Absenteeism\%} = \frac{\text{ChronicAbsCount}}{\text{EligibleEnroll}} \times 100$$

Panel Construction and Merge Keys

The dataset was constructed by merging multiple sources using either `SchoolCode` or a truncated `CDSCode`. After cleaning and harmonization, the final unified dataset includes 1,011 schools present in the 2021–2022 academic year. Schools that appeared multiple times in the BTB program rosters, due to separate listings for before- and after-school programs, were aggregated into a single school level binary variable. A school was coded as participating in BTB if it appeared in any program roster.

4.3 Unit of Analysis

The unit of analysis is a school in LAUSD between the 2021–2022 academic year.

4.4 Key Outcome Variables

Outcome variables represent observable indicators that may shift in response to treatment status, which is determined by whether a school's running variable crosses a policy cutoff.

The analysis aims to detect any discontinuities in these outcomes at the threshold, assuming the relationship between the running variable and the outcome would otherwise remain smooth.

This study focuses on school level academic achievement, student attendance, and access to supplemental programs. These categories reflect both the goals of California's education funding policies and the pathways through which funding and services may influence student outcomes.

Academic achievement is captured using CAASPP (California Assessment of Student Performance and Progress) data, including both proficiency rates and average scale scores in English Language Arts (ELA) and Mathematics. Attendance is measured by the chronic absenteeism rate, defined as the percentage of students who missed 10 percent or more of instructional days for any reason. Program access is measured using a binary indicator for participation in the Beyond the Bell (BTB) before- and after-school program, a key LAUSD offering designed to expand learning time.

These outcomes were selected based on their relevance to state and district policy objectives, their importance for equity and accountability, and their availability in publicly accessible administrative datasets. Table 4.2 lists and describes all outcome variables included in the analysis.

4.5 Running Variables

In a regression discontinuity design, the running variable is a continuous measure that determines whether a unit receives treatment. Units with values above a specific cutoff are assigned to the treatment group, while those below are assigned to the control group. This cutoff generates local variation that can be treated as quasi-random, allowing researchers to estimate causal effects. The key identifying assumption is that, in the absence of treatment, the relationship between the running variable and the outcome would be smooth at the threshold.

In the present study, two types of continuous school level measures serve as running variables: the percentage of students eligible for Free and Reduced-Price Meals (FRPM) and the percentage of unduplicated pupils, which includes students who are low income, English learners, or foster youth. These measures are used by California education policy to determine eligibility for critical supports, including Title I funding and supplemental LCFF grants. Table 4.3 summarizes the specific thresholds and their corresponding policy rules, which create discontinuities leveraged for causal identification.

4.6 Covariates

Although regression discontinuity designs rely on the local variation created by the running variable cutoff, adding covariates can improve the precision of treatment effect estimates and support validity checks. In this study, covariates include continuous demographic measures such as racial and ethnic composition and total enrollment, as well as categorical indicators for school type, including charter status and DASS designation. These variables are not expected to shift sharply at the threshold, but they help account for residual variation in outcomes. Table 4.4 lists and describes the covariates included in the analysis. Controlling for these characteristics helps ensure more stable estimates and allows for additional tests of balance across the cutoff.

Outcome Variable	Description
% Met Above Standard – ELA	Percent of students scoring at or above the proficiency standard in English Language Arts on the CAASPP assessment.
% Met Above Standard – Math	Percent of students scoring at or above the proficiency standard in Mathematics on the CAASPP assessment.
Chronic Absenteeism Rate	Share of enrolled students who were absent for 10% or more of the school year for any reason.
BTB Participation	Binary indicator equal to 1 if the school offered any Beyond the Bell (BTB) before- or after-school programming.
% Not Met Standard – Math	Percent of students who did not meet the standard in Mathematics on the CAASPP assessment.
Scale Score – ELA	Average English Language Arts scale score on the CAASPP assessment for the school.
Scale Score – Math	Average Mathematics scale score on the CAASPP assessment for the school.

Table 4.2: Outcome Variables Used in the GP-RDD Models

Running Variable	Cutoff	Policy Use
% FRPM	35%	Minimum Title I eligibility [Edu21b]
% FRPM	75%	First Priority Title I allocation [Edu21b]
% Unduplicated Pupils	55%	School Site Council planning requirement [Inf21]
% Unduplicated Pupils	75%	LCFF Concentration Grant funding intensifies [Edu21a]

Table 4.3: Policy Based Cutoffs for Common school level Running Variables in California

Covariates	Description
pct_hispanic	Percentage of enrolled students identified as Hispanic or Latino
pct_black	Percentage of enrolled students identified as Black or African American
pct_white	Percentage of enrolled students identified as White
pct_asian	Percentage of enrolled students identified as Asian
pct_two_or_more	Percentage of students identifying with two or more races
pct_other	Percentage of students identified with another racial/ethnic group
total_enroll	Total number of students enrolled at the school
Charter.School	Indicator for whether the school is a charter school (1 = charter, 0 = traditional)
DASS	Indicator for participation in the Dashboard Alternative School Status (DASS) program

Table 4.4: Covariates used in balance tests and adjusted models

CHAPTER 5

Methodology

5.1 Introduction and Research Design

This chapter outlines the causal identification strategy used to estimate the effect of program eligibility thresholds on educational outcomes in LAUSD. We use these funding thresholds to estimate local causal effects using Regression Discontinuity Design (RDD).

5.1.1 Causal Research Question

The central research question is: *What is the causal effect of program funding eligibility on chronic absenteeism, beyond the bell participation, and CAASPP scoring categories in LAUSD schools?*

Randomized experiments are difficult in educational policy, but when eligibility is governed by cutoffs, RDD provides a natural quasi-experiment. Schools just above and just below a threshold are similar in all observable and unobservable characteristics, except for program access, allowing us to estimate the treatment effect (τ) as the causal impact of eligibility at that threshold. In this study, we use the following thresholds: FRPM 35% and 75% and UPP 55% and 75% cut off thresholds.

In a **RDD**, the assignment of treatment is determined by whether the running variable exceeds a known cutoff.

$$D_i = \begin{cases} 1 & \text{if } X_i \geq c \\ 0 & \text{if } X_i < c \end{cases}$$

Three assumptions underpin the validity of RDD:

1. **Continuity of the potential outcome:** In the absence of treatment, the conditional expectation of Y_i is continuous at the threshold:

$$\lim_{x \uparrow c} \mathbb{E}[Y_i(0)] = \lim_{x \downarrow c} \mathbb{E}[Y_i(0)]$$

2. **No manipulation of the running variable:** Schools cannot precisely sort around the cutoff. FRPM eligibility is based on household income forms or categorical qualification, which are externally determined and thus unlikely to be manipulated.

3. **Local randomization:** Within a narrow bandwidth around the cutoff, schools are as if randomly assigned to treatment and control conditions.

5.2 Classical RDD Estimation with `rdrobust`

When called with its default arguments, `rdrobust` implements a conventional *sharp* regression discontinuity (RD) design that estimates a level discontinuity at the cutoff c (which defaults to 0 unless otherwise specified). The routine first selects bandwidths by invoking the companion selector `rdbwselect`. This uses local curvature estimates on each side of the cutoff to choose main bandwidths h_- and h_+ and corresponding bias-correction bandwidths b_- and b_+. If curvature differs across sides, these widths may differ as well. Within those bandwidths, `rdrobust` fits separate local-linear regressions ($p = 1$) on the left and right using the triangular kernel $K(u) = \big(1 - |u|\big)\mathbf{1}\big\{|u| < 1\big\}$, which down-weights observations farther from the threshold.

5.3 Bayesian Gaussian Process RDD Estimation with gpss

Local polynomial estimators such as `rdrobust` achieve efficiency in large, data-rich designs because the bias they introduce at the boundary is both small and analytically correctable. In modest samples, however, that same structure becomes fragile. A single polynomial must extrapolate into the sparsest part of the running-variable support, so its fitted residual variance shrinks precisely where variance should explode. The result is a variance estimator that is *too optimistic*, producing confidence intervals that include the true treatment effect *less frequently than the nominal 95 % rate.* Hazlett *et al.*'s simulations make this concrete: with $n = 250$ observations, 's nominal 95 % intervals cover only 80-84 % of the time and its RMSE is roughly 40 % larger than the ground truth. Those failures stem from the estimator's dependence on a single functional form and a bandwidth h chosen to minimize mean-squared error yet blind to the full distribution of functions that fit the data equally well.

A **GP-RDD** addresses each of these weaknesses by replacing the fixed polynomial with a *distribution over functions.* The only prior structure is the covariance kernel $k_\theta\big((R_i, \mathbf{X}_i), (R_j, \mathbf{X}_j)\big)$. Consequently, posterior variance *inflates* as $\|R - c\|$ grows, yielding intervals that remain calibrated even when observations are sparse or the conditional-expectation function bends sharply.

In our implementation, covariates are appended to the running variable as inputs to the GP kernel. The GP mean remains zero, so no linear trend is imposed or "double counted." This kernel-based inclusion allows flexible, potentially nonlinear adjustment for W. We therefore interpret any loss of significance after adding W as evidence that the unadjusted effect was partly explained by observed characteristics and/or that power is limited locally around the threshold.

The GP-RDD relies on the standard RD identification assumptions of continuity and no precise manipulation. On the modeling side, Gaussian processes define a large, flexible function space via the kernel, so posterior variance naturally inflates at the data's edge, "appropriately reflecting increased uncertainty … in regions with little or no data" [CKH2X]. Estimation further assumes Gaussian, homoskedastic errors, since σ^2 is selected by marginal likelihood under that model. Thus, rather than residual checks, model assessment rests on whether predictive intervals widen with sparser support.

CHAPTER 6

Results

6.1 RD Robust Results

Table 6.1 presents the local linear RDD estimates for selected outcomes at the 35% and 75% FRPM thresholds, estimated using the `rdrobust` package. At the 35% FRPM threshold, the estimated treatment effect on chronic absenteeism is 6.48 percentage points (SE = 4.22), with a 95% confidence interval of [-1.78, 14.74], which is not statistically significant ($p = 0.124$). The strongest effects at this threshold are observed in academic outcomes: the percentage of students meeting the Math proficiency standard decreases by 17.19 percentage points (SE = 9.42, $p = 0.068$), while the percentage not meeting the Math standard increases by 14.56 percentage points (SE = 7.74, $p = 0.060$). Similarly, a marginally significant increase of 7.93 percentage points is observed in the share of students not meeting the ELA standard ($p = 0.054$).

At the 75% FRPM threshold, results are weaker across all outcomes. Chronic absenteeism shows a small, statistically insignificant increase of 1.49 percentage points (SE = 4.69, $p = 0.751$). The percentage of students meeting the ELA standard increases by 8.53 percentage points (SE = 4.48), but the confidence interval includes zero and the result is marginal ($p = 0.057$). Other academic outcomes such as Math proficiency and scale scores also yield imprecise estimates with wide confidence intervals and p-values well above conventional significance levels.

At the 35% FRPM cutoff, the direction of the effects raises important concerns. For both ELA and Math proficiency, point estimates are negative, meaning that schools just crossing into eligibility exhibit lower proficiency rates than those just below. For illustration using Table 6.1, at the 35% FRPM cutoff the classical RDD estimate for % Met Above *Math* is -17.19 percentage points (SE $= 9.42$, $p = 0.068$). For % Met Above, *ELA* it is -9.42 percentage points (SE $= 6.42$, $p = 0.143$). Directionally, schools just crossing into eligibility exhibit lower proficiency than those just below, but only the Math estimate approaches conventional significance in the classical specification.

Table 6.1: Classical RDD Results with Highlighted Significance

Outcome	Running Variable	Cutoff	τ	SE	CI Lower	CI Upper	p-value
chronic_absenteeism	`frpm_percent`	35	6.478	4.216	-1.784	14.741	0.124
btb	`frpm_percent`	35	-0.127	0.162	-0.445	0.190	0.432
avg_pct_met_above_ELA	`frpm_percent`	35	-9.420	6.423	-22.009	3.170	0.143
avg_pct_met_above_Math	`frpm_percent`	35	-17.191	9.421	-35.655	1.274	0.068
avg_pct_not_met_ELA	`frpm_percent`	35	7.928	4.115	-0.137	15.993	0.054
avg_pct_not_met_Math	`frpm_percent`	35	14.556	7.738	-0.609	29.722	0.060
avg_scale_score_ELA	`frpm_percent`	35	-19.381	23.225	-64.901	26.138	0.404
avg_scale_score_Math	`frpm_percent`	35	-43.856	27.243	-97.251	9.540	0.107
chronic_absenteeism	`frpm_percent`	75	1.488	4.687	-7.698	10.674	0.751
btb	`frpm_percent`	75	-0.319	0.173	-0.658	0.020	0.065
avg_pct_met_above_ELA	`frpm_percent`	75	8.530	4.482	-0.254	17.315	0.057
avg_pct_met_above_Math	`frpm_percent`	75	6.781	6.247	-5.462	19.025	0.278
avg_pct_not_met_ELA	`frpm_percent`	75	-7.452	3.894	-15.083	0.179	0.056
avg_pct_not_met_Math	`frpm_percent`	75	-8.804	7.020	-22.562	4.954	0.210
avg_scale_score_ELA	`frpm_percent`	75	25.373	16.088	-6.159	56.904	0.115
avg_scale_score_Math	`frpm_percent`	75	19.637	11.306	-2.523	41.797	0.082
chronic_absenteeism	`undup_pct`	55	-0.188	3.485	-7.019	6.643	0.957
btb	`undup_pct`	55	0.110	0.224	-0.330	0.550	0.623
avg_pct_met_above_ELA	`undup_pct`	55	8.486	3.543	1.542	15.429	0.017
avg_pct_met_above_Math	`undup_pct`	55	7.368	6.005	-4.402	19.137	0.220
avg_pct_not_met_ELA	`undup_pct`	55	-0.657	2.825	-6.194	4.880	0.816
avg_pct_not_met_Math	`undup_pct`	55	-5.347	4.977	-15.102	4.407	0.283
avg_scale_score_ELA	`undup_pct`	55	7.918	29.685	-50.263	66.100	0.790
avg_scale_score_Math	`undup_pct`	55	8.085	19.270	-29.683	45.853	0.675
chronic_absenteeism	`undup_pct`	75	-1.684	4.928	-11.342	7.974	0.733

Outcome	Running Variable	Cutoff	τ	SE	CI Lower	CI Upper	p-value
btb	`undup_pct`	75	-0.007	0.174	-0.348	0.334	0.970
avg_pct_met_above_ELA	`undup_pct`	75	6.400	3.953	-1.348	14.148	0.105
avg_pct_met_above_Math	`undup_pct`	75	9.840	5.053	-0.063	19.744	0.051
avg_pct_not_met_ELA	`undup_pct`	75	-2.805	3.709	-10.075	4.466	0.450
avg_pct_not_met_Math	`undup_pct`	75	-7.134	5.851	-18.603	4.334	0.223
avg_scale_score_ELA	`undup_pct`	75	1.452	20.822	-39.359	42.263	0.944
avg_scale_score_Math	`undup_pct`	75	7.400	12.853	-17.791	32.591	0.565

6.2 Bayesian GP-RDD Results

Figures 6.1, 6.2, and 6.3 plot the GP-RDD posterior means and 95% credible intervals for selected academic outcomes around the 35% FRPM threshold. At this cutoff, the Bayesian model identifies statistically significant treatment effects in multiple areas. The estimated effect on the percentage of students meeting Math proficiency standards is -20.57 percentage points (SE = 7.39, 95% CI: $[-35.05, -6.09]$, $p = 0.005$), while the percentage not meeting Math standards increases by 17.33 percentage points (SE = 6.60, 95% CI: $[4.40, 30.27]$, $p = 0.009$). Additionally, the share of students meeting ELA proficiency standards declines by 13.26 percentage points (SE = 6.73, 95% CI: $[-26.44, -0.07]$, $p = 0.049$). These estimates are statistically significant and substantively large, suggesting that crossing the 35% FRPM threshold is associated with worsened academic performance in both Math and ELA.

In contrast, the estimated effect on chronic absenteeism at the 35% threshold is not statistically significant, with a posterior mean of 7.87 percentage points (SE = 5.74, 95% CI: $[-3.38, 19.11]$, $p = 0.171$). Results for the 75% FRPM threshold are uniformly non-significant across all outcomes, including absenteeism, where the estimated effect is -0.07 percentage points (SE = 6.06, 95% CI: $[-11.95, 11.80]$, $p = 0.990$). These findings reinforce that the most robust effects are concentrated at the 35% FRPM cutoff and are limited to academic achievement, rather than attendance.

In addition to statistical significance, the GP-RDD estimates also reveal troubling directions for the treatment effects. The negative coefficients for % Met Above Standard in both Math (18.52 pp, p = 0.013) and ELA (14.90 pp, p = 0.027) indicate that crossing the 35% FRPM threshold is associated with a substantial drop in proficiency. Conversely, the positive coefficients for % Not Met Standard in Math (+14.29 pp, p = 0.038) and ELA (+11.04 pp, p = 0.049) show an increase in low performance rates. Interpreted causally, these results suggest that schools just above the 35% cutoff perform worse academically than those just below, contradicting the expected benefits of eligibility. This negative direction is consistent across unadjusted models, though attenuated or lost after covariate adjustment.

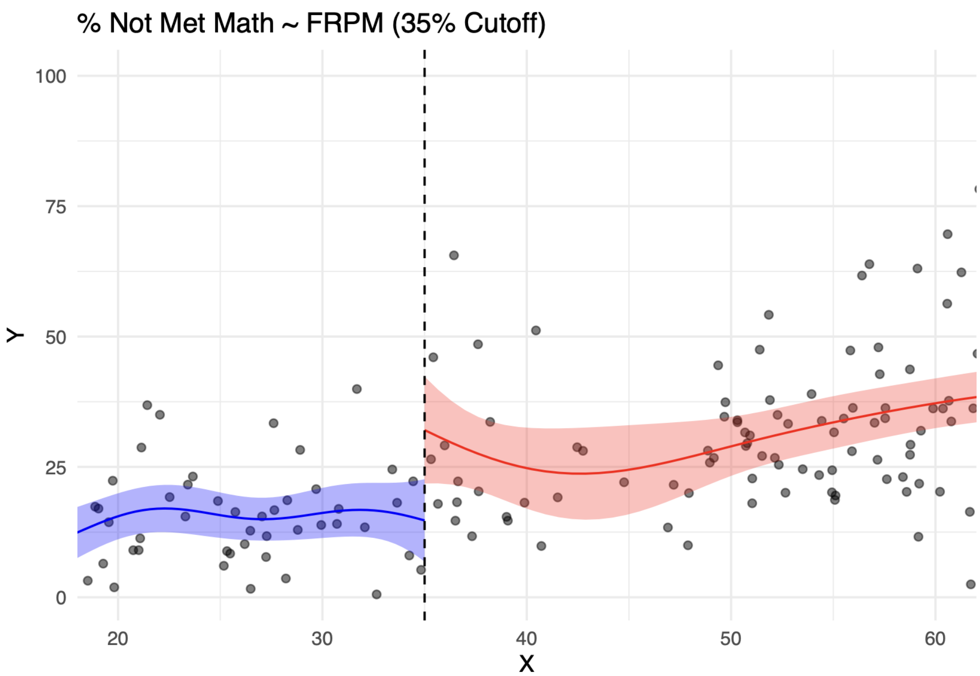

Figure 6.1: Effect of FRPM eligibility (35% cutoff) on % Not Met ELA performance.

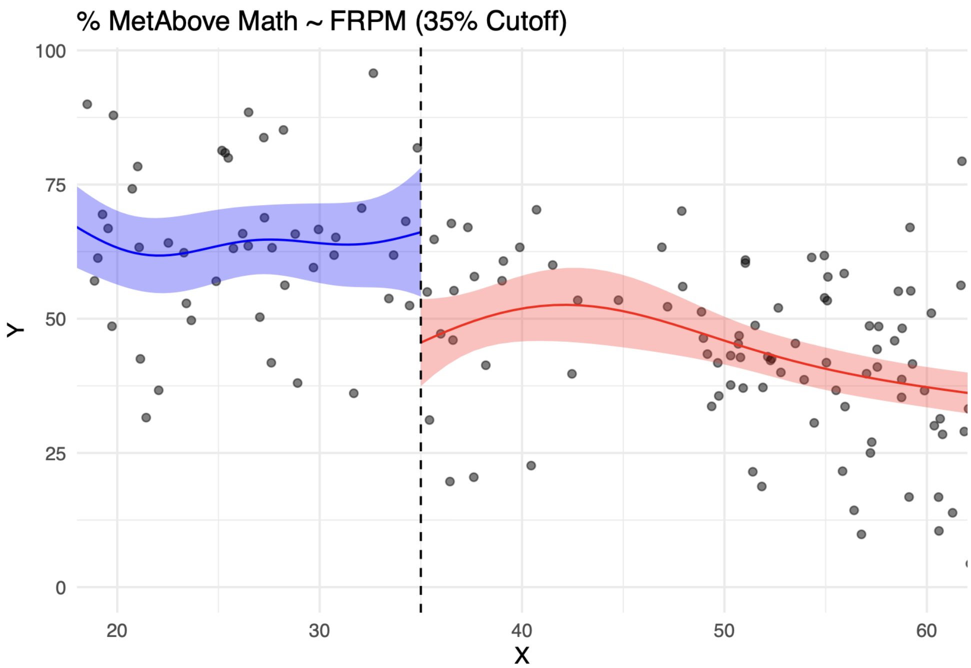

Figure 6.2: Effect of FRPM eligibility (35% cutoff) on % Met Above Math performance.

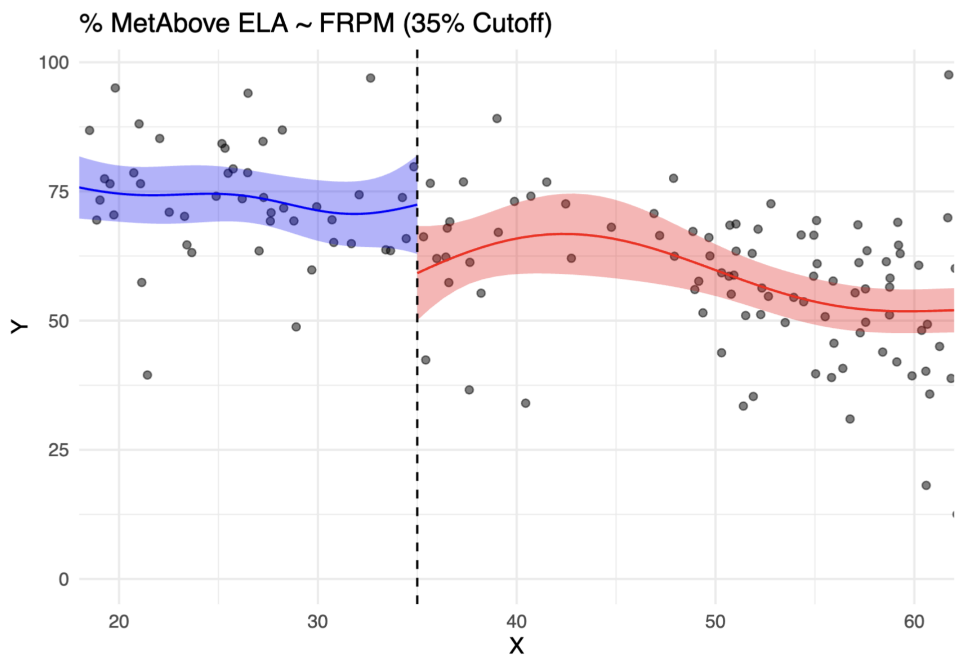

Figure 6.3: Effect of FRPM eligibility (35% cutoff) on % Met Above ELA performance.

Table 6.2: Gaussian Process RDD Results (Significant Rows Highlighted)

Model	τ	SE	CI Lower	CI Upper	p-value
% MetAbove ELA FRPM (35%)	-13.26	6.73	-26.44	-0.07	0.0488
% MetAbove ELA FRPM (75%)	4.23	5.99	-7.51	15.98	0.4800
% MetAbove ELA Undup (55%)	-1.72	7.64	-16.69	13.26	0.8223
% MetAbove ELA Undup (75%)	-1.76	6.43	-14.36	10.84	0.7842
% MetAbove Math FRPM (35%)	-20.57	7.39	-35.05	-6.09	0.0054
% MetAbove Math FRPM (75%)	4.42	5.59	-6.53	15.37	0.4294
% MetAbove Math Undup (55%)	-2.11	7.40	-16.62	12.40	0.7758
% MetAbove Math Undup (75%)	3.74	5.99	-8.01	15.49	0.5330
Chronic Absenteeism FRPM (35%)	7.87	5.74	-3.38	19.11	0.1706
Chronic Absenteeism FRPM (75%)	-0.07	6.06	-11.95	11.80	0.9903
BTB FRPM (35%)	-0.23	0.22	-0.67	0.21	0.3029
BTB FRPM (75%)	-0.16	0.21	-0.57	0.25	0.4445
Chronic Absenteeism Undup (55%)	3.06	7.58	-11.80	17.93	0.6866
Chronic Absenteeism Undup (75%)	-1.67	6.68	-14.75	11.42	0.8027
% Not Met Math FRPM (35%)	17.33	6.60	4.40	30.27	0.0086
% Not Met Math FRPM (75%)	-6.40	6.78	-19.69	6.88	0.3449
% Not Met Math Undup (55%)	3.28	8.50	-13.38	19.94	0.6995
% Not Met Math Undup (75%)	-3.08	7.41	-17.60	11.45	0.6782
Scale ELA FRPM (35%)	-26.29	26.46	-78.16	25.57	0.3204
Scale ELA FRPM (75%)	-3.88	26.34	-55.51	47.75	0.8830
Scale ELA Undup (55%)	-9.84	32.98	-74.47	54.79	0.7654
Scale ELA Undup (75%)	-25.92	28.64	-82.06	30.22	0.3655
Scale Math FRPM (35%)	-42.42	22.45	-86.42	1.58	0.0588

Model		τ	SE	CI Lower	CI Upper	p-value
Scale Math	FRPM (75%)	-0.75	19.36	-38.69	37.20	0.9692
Scale Math	Undup (55%)	-7.18	24.92	-56.02	41.65	0.7731
Scale Math	Undup (75%)	-17.78	20.98	-58.89	23.34	0.3967

6.3 Classical vs. Gaussian Process Comparison

This section compares the treatment effect estimates obtained from the classical local linear RDD framework to those produced by the GP-RDD. While both approaches aim to recover local average treatment effects at sharp eligibility cutoffs, they differ in foundational assumptions and in how uncertainty is quantified. The classical RDD relies on nonparametric frequentist estimation with fixed bandwidths and polynomial specifications, whereas GP-RDD adopts a fully Bayesian framework that models the outcome surface flexibly using a Gaussian process prior and quantifies uncertainty through posterior distributions.

Across both methods, results are broadly consistent in direction and magnitude for the strongest effects. For instance, at the FRPM 35% cutoff, the estimated effect on `avg_pct_met_above_Math` is $\hat{\tau} = -17.19$ (SE = 9.42, $p = 0.068$) under the classical RDD, and $\hat{\tau} = -20.57$ (SE = 7.39, $p = 0.005$) under GP-RDD. The Bayesian model yields a narrower credible interval and crosses the conventional significance threshold, while the classical estimate is marginal. Similarly, the estimated effect on `avg_pct_not_met_Math` at the same cutoff is positive and sizable in both models: 14.56 (SE = 7.74, $p = 0.060$) classically, and 17.33 (SE = 6.60, $p = 0.009$) via GP-RDD.

In some cases, GP-RDD identifies statistically significant effects where the classical model does not. For `avg_pct_met_above_ELA` at the FRPM 35% cutoff, the classical RDD yields $\hat{\tau} = -9.42$ (SE = 6.42, $p = 0.143$), while the GP-RDD estimate is $\hat{\tau} = -13.26$ (SE = 6.73, $p = 0.049$). Though the direction is consistent, only the Bayesian credible interval excludes zero. This pattern recurs for several outcomes near the boundary of significance, where the GP-RDD's posterior distribution places greater mass away from zero due to smoother function estimation across the running variable.

Conversely, the two methods diverge slightly in cases with weak or null effects. For `chronic_absenteeism` at the FRPM 75% threshold, the classical estimate is 1.49 (SE = 4.69, $p = 0.751$), while GP-RDD estimates -0.07 (SE = 6.06, $p = 0.990$). Though both are statistically indistinguishable from zero, the direction differs slightly, reflecting differences in local fit due to bandwidth selection versus GP kernel smoothing.

6.4 Covariate Balance Tests

To support the validity of the regression discontinuity design (RDD), we conduct a series of covariate balance tests around each threshold. The goal of these tests is to verify the assumption that schools just above and below the eligibility cutoffs are similar in observable characteristics, thereby justifying the interpretation of estimated effects as causal.

We implement Gaussian Process RDD based balance tests on both continuous and binary covariates, examining potential discontinuities at each cutoff for both running variables. Specifically, we test for smoothness in:

- Racial and ethnic composition: percent Hispanic, Black, White, Asian, multiracial, and other
- Total enrollment
- Binary characteristics such as charter school status and DASS classification

Each test produces an estimate $\hat{\tau}$, a standard error, a 95% credible interval, and a p-value. A statistically insignificant discontinuity supports the assumption of covariate continuity at the threshold. To ensure robustness, we repeat this procedure at four cutoffs: 35% and 75% for FRPM, and 55% and 75% for UPP.

Table 6.3: Balance Test Results from Gaussian Process RDD (Significant Rows Highlighted)

Variable	Cutoff	τ	SE	CI Lower	CI Upper	p-value
pct_hispanic	FRPM (35%)	16.40	6.82	3.04	29.78	0.0162
pct_black	FRPM (35%)	3.49	5.10	-6.50	13.48	0.4934
pct_white	FRPM (35%)	-11.70	6.05	-23.55	0.17	0.0534
pct_asian	FRPM (35%)	-7.62	4.42	-16.27	1.04	0.0848
pct_two_or_more	FRPM (35%)	0.64	1.77	-2.83	4.10	0.7189
pct_other	FRPM (35%)	-0.55	4.83	-10.01	8.92	0.9096
total_enroll	FRPM (35%)	-21.00	304.53	-617.89	575.86	0.9450
Charter.School	FRPM (35%)	0.15	0.24	-0.32	0.63	0.5308
DASS	FRPM (35%)	0.01	0.08	-0.16	0.17	0.9306
pct_hispanic	FRPM (75%)	6.27	7.28	-8.00	20.53	0.3893
pct_black	FRPM (75%)	2.51	6.02	-9.29	14.31	0.6767
pct_white	FRPM (75%)	-6.81	3.99	-14.64	1.02	0.0884
pct_asian	FRPM (75%)	0.01	2.23	-4.37	4.39	0.9966
pct_two_or_more	FRPM (75%)	0.03	0.86	-1.65	1.71	0.9726
pct_other	FRPM (75%)	-1.94	2.26	-6.38	2.50	0.3922
total_enroll	FRPM (75%)	1030.00	279.22	481.86	1576.39	0.0002
Charter.School	FRPM (75%)	0.15	0.19	-0.22	0.52	0.4338
DASS	FRPM (75%)	0.03	0.10	-0.17	0.22	0.7931
pct_hispanic	Undup (55%)	16.50	8.71	-0.61	33.52	0.0587
pct_black	Undup (55%)	-1.58	7.21	-15.72	12.56	0.8266
pct_white	Undup (55%)	-18.50	5.01	-28.35	-8.70	0.0002
pct_asian	Undup (55%)	5.98	3.22	-0.33	12.28	0.0631
pct_two_or_more	Undup (55%)	-0.56	1.22	-2.94	1.82	0.6456

Variable	Cutoff	τ	SE	CI Lower	CI Upper	p-value
pct_other	Undup (55%)	-2.46	3.00	-8.33	3.41	0.4118
total_enroll	Undup (55%)	-164.00	365.77	-880.81	552.99	0.6541
Charter.School	Undup (55%)	-0.08	0.25	-0.56	0.41	0.7507
DASS	Undup (55%)	-0.01	0.12	-0.24	0.23	0.9581
pct_hispanic	Undup (75%)	-5.81	7.56	-20.62	9.00	0.4420
pct_black	Undup (75%)	1.51	6.55	-11.33	14.35	0.8178
pct_white	Undup (75%)	5.55	3.97	-2.23	13.33	0.1623
pct_asian	Undup (75%)	2.89	2.37	-1.76	7.54	0.2232
pct_two_or_more	Undup (75%)	-1.04	0.84	-2.69	0.61	0.2177
pct_other	Undup (75%)	-3.50	2.31	-8.02	1.02	0.1287
total_enroll	Undup (75%)	-505.00	301.29	-1095.53	85.51	0.0937
Charter.School	Undup (75%)	0.05	0.21	-0.35	0.45	0.8039
DASS	Undup (75%)	0.06	0.11	-0.15	0.28	0.5619

6.5 Covariate-Adjusted RDD Models

We estimate covariate-adjusted models that include a predictive set of pre-treatment covariates W to improve precision. In GP-RDD, these covariates enter *through the kernel*, allowing the GP to flexibly partial out systematic variation associated with W without imposing a linear functional form. Balance tests serve as diagnostics rather than as a gating criterion for inclusion. These adjustments improve precision and partially account for residual differences between treatment and control schools, though they do not resolve violations of continuity in potential outcomes.

The classical adjusted model takes the form:

$$Y_i = \alpha + \tau D_i + \beta(X_i - c) + \gamma^\top W_i + \varepsilon_i$$

where:

- Y_i is the outcome (e.g., absenteeism or test score)
- $D_i = \mathbb{1}(X_i \geq c)$ is an indicator for crossing the threshold
- X_i is the running variable (centered at cutoff c)
- W_i is a vector of imbalanced pre-treatment covariates
- τ is the treatment effect of interest

In adjusted specifications, we include a pre-treatment covariate set W to improve precision rather than to correct detected discontinuities. In the GP-RDD, W enters through the covariance kernel so that similarity is assessed jointly in the running variable and W. Balance tests are reported separately in 6.3 and serve as diagnostics, not as a gating criterion for inclusion.

6.6 Comparison of Adjusted and Unadjusted GP-RDD Models

To assess the robustness of the estimated treatment effects, we compare results from standard GP-RDD models to those from covariate-adjusted GP-RDD specifications. The adjusted models include a set of school level covariates identified as imbalanced in the covariate balance tests: percent Hispanic, percent White, percent multiracial, Charter School status, and DASS status. These covariates are incorporated into the Gaussian process via the kernel function, allowing the GP to account for residual variation after adjusting for observable characteristics in a flexible way.

At the 35% FRPM cutoff, several effects that are statistically significant in the unadjusted GP-RDD become non-significant once covariates enter the kernel (see 6.4, e.g., % Met Above ELA and % Met Above Math). We interpret this change in two ways. First, kernel-based covariate adjustment can *remove bias* if observed composition near the cutoff explains part of the discontinuity. Second, adding covariates increases the input dimensionality the GP must learn locally. This tempers claims that eligibility alone causes the observed changes at the threshold.

Because several GP-RDD estimates lose significance after kernel-based covariate adjustment, we regard those discontinuities as *sensitive* to observed composition near the cutoff and/or to limited local sample size. The direction of effects is often stable, but once observed characteristics are flexibly controlled, interval estimates include zero. These outcomes we present the evidence as *suggestive rather than definitive*.

Table 6.4: GP-RDD and RDD Estimates Across Specifications (Significance Highlighted in p-value Column)

Outcome & Cutoff	**RDD**					**GP-RDD (no covs)**					**GP-RDD (covariate-adjusted)**				
	τ	SE	CI Lower	CI Upper	p-value	τ	SE	CI Lower	CI Upper	p-value	τ	SE	CI Lower	CI Upper	p-value
avg_pct_met_above_ELA FRPM (35%)	-12.85	6.00	-24.61	-1.09	0.0323	-14.90	6.74	-28.12	-1.69	0.0271	-5.15	9.86	-24.47	14.17	0.6013
avg_pct_met_above_Math FRPM (35%)	-15.07	8.70	-32.12	1.98	0.0832	-18.52	7.43	-33.07	-3.96	0.0126	-9.30	10.11	-29.11	10.51	0.3576
avg_pct_not_met_ELA FRPM (35%)	9.38	4.54	0.48	18.27	0.0389	11.04	5.60	0.07	22.00	0.0486	4.28	8.88	-13.12	21.69	0.6295
avg_pct_not_met_ELA FRPM (75%)	-11.12	4.75	-20.44	-1.81	0.0193	-2.40	5.94	-14.04	9.24	0.6861	-5.08	10.22	-25.11	14.95	0.6192
avg_pct_not_met_Math FRPM (35%)	11.28	5.87	-0.23	22.78	0.0547	14.29	6.88	0.80	27.78	0.0379	4.48	10.54	-16.18	25.14	0.6707

Across adjusted and unadjusted specifications, signs and magnitudes are similar and covariate adjustment typically reduces standard errors. However, in a few cases at the 35% FRPM cutoff, effects that are statistically significant in unadjusted GP-RDD become non-significant once covariates enter the kernel (see 6.4).

CHAPTER 7

Discussion

7.1 Interpretation of Main Findings

Table 6.4 reports three specifications: a classical local–linear RDD, a Bayesian GP-RDD without covariates, and a GP-RDD with kernel-based covariate adjustment. By design, treatment occurs when crossing from below to above each cutoff. Negative τ on "% Met Above" indicates lower proficiency for newly eligible schools. Positive τ on "% Not Met" indicates worse performance. These sign conventions orient each estimate to "what happens when a school becomes eligible," so negative proficiency and positive non-proficiency values indicate locally worse outcomes just above the threshold.

FRPM 35%: ELA proficiency. All three models point to a decline at the cutoff. RDD estimates -12.85 percentage points (pp) with $p = 0.0323$, GP-RDD estimates -14.90 pp with $p = 0.0271$, and the covariate-adjusted GP-RDD estimate attenuates to -5.15 pp and is not significant ($p = 0.6013$) (Table 6.4). Crossing into eligibility at 35% is associated with a lower share of students meeting ELA standards just above the cutoff, and once observed composition is modeled in the GP kernel the remaining difference is small and statistically indistinguishable from zero.

FRPM 35%: Math proficiency. For % Met Above Math, RDD is -15.07 pp and marginal ($p = 0.0832$). GP-RDD is -18.52 pp and significant ($p = 0.0126$). With kernel covariates, the estimate is -9.30 pp and not significant ($p = 0.3576$) (Table 6.4). Crossing into eligibility at 35% is associated with a notably lower share of students meeting Math standards for schools just above the cutoff, but this apparent decline becomes smaller and not distinguishable from zero after covariate adjustment in the GP model.

FRPM 35%: ELA non-proficiency. For % Not Met ELA, RDD indicates +9.38 pp ($p = 0.0389$), GP-RDD indicates +11.04 pp ($p = 0.0486$), and the covariate-adjusted GP-RDD estimate is +4.28 pp and not significant ($p = 0.6295$) (Table 6.4). Immediately above 35%, a larger share of students do not meet ELA standards relative to just-below schools, yet this increase is not distinguishable from zero once covariates are added.

FRPM 35%: Math non-proficiency. For % Not Met Math, RDD is +11.28 pp and marginal ($p = 0.0547$). GP-RDD is +14.29 pp and significant ($p = 0.0379$). With kernel covariates, the estimate is +4.48 pp and not significant ($p = 0.6707$) (Table 6.4). Immediately above 35%, a larger share of students do not meet Math standards, but this increase disappears after covariate adjustment.

FRPM 75%: ELA non-proficiency. At the higher 75% cutoff, RDD shows a decrease of -11.12 pp ($p = 0.0193$), while both GP specifications are small and not significant (GP-RDD -2.40 pp, $p = 0.6861$; covariate-adjusted GP-RDD -5.08 pp, $p = 0.6192$) (Table 6.4). Classical RDD implies fewer students fail to meet ELA standards just above 75%, but the GP specifications do not detect a difference, so this isolated RDD finding should be treated cautiously.

Composition at the cutoff. Balance tests (Chapter 6) indicate observable differences around some thresholds. At FRPM 35%, schools just above the cutoff have a higher share of Hispanic students (approximately +16.4 pp, $p \approx 0.016$) and a lower share of White students (about -11.7 pp, $p \approx 0.053$); Asian share is borderline (about -7.6 pp, $p \approx 0.085$). At FRPM 75%, schools above the cutoff are much larger on average (roughly +1030 students,

$p \approx 0.0002$). These patterns are consistent with the changes seen in the covariate-adjusted GP-RDD and with wider local uncertainty.

Summary. At 35% FRPM, proficiency falls and non-proficiency rises just above the cutoff across models. The unadjusted GP-RDD detects these effects most clearly. Once covariates are included in the kernel, estimates shrink and are not statistically significant, which is consistent with observable composition near the cutoff accounting for part of the discontinuity and with wider uncertainty in the local window. At 75% FRPM, effects are generally null in the GP models, with one classical RDD exception in ELA non-proficiency. Taken together, gaining eligibility at 35% coincides locally with worse academic performance, but the apparent differences are sensitive to observed composition and weaken once those characteristics are modeled.

7.2 Policy Implications

For the 35% FRPM threshold, the combined evidence suggests that schools just above the cutoff perform worse on achievement than schools just below, but these effects are not robust to covariate adjustment in the GP framework (Table 6.4). The direction and attenuation indicate either negative selection just above the cutoff or factors correlated with eligibility that the local designs do not capture. These are local effects at the threshold and speak to eligibility rather than confirmed receipt or intensity of services.

For the 75% FRPM threshold, there is no consistent evidence of effects. Classical RDD detects an improvement for ELA non-proficiency that the GP models do not replicate, so policy conclusions at this boundary should be cautious.

Attendance and participation outcomes are not included in Table 6.4 and are therefore not discussed here. The safest conclusion is that threshold-based eligibility in this period does not consistently improve achievement near the cutoffs and may coincide with worse outcomes at FRPM 35 in unadjusted analyses. Any decision to expand, contract, or retarget programs should be paired with follow-up work that tracks intensity of services, lags between eligibility and implementation, and outcomes over multiple years.

CHAPTER 8

Limitations

8.1 Methodological Limitations

First, while the primary design is modeled as a sharp regression discontinuity, the reality of program implementation introduces an element of fuzzy compliance. In particular, BTB participation does not deterministically follow FRPM thresholds; some schools below the threshold participate while others above it do not. This partial compliance weakens the first stage relationship and implies that the estimated local average treatment effect (LATE) should be interpreted conditional on actual treatment uptake.

Second, there remains the possibility of unobserved confounding or omitted variable bias near the threshold. Although the continuity assumption ensures that observed and unobserved covariates are smoothly distributed around the cutoff, certain unmeasured factors such as school leadership quality, neighborhood safety, or local funding initiatives, which could still correlate with both FRPM rates and absenteeism outcomes.

Third, a fundamental limitation of RDD is its inherently local nature. Interpreting the direction of effects assumes that model specification and identification assumptions hold. While the consistency of negative estimates across methods lends credibility, the attenuation of some effects after covariate adjustment suggests that caution is warranted in drawing strong causal conclusions about the sign. It remains possible that unobserved factors correlated with crossing the threshold could contribute to these patterns. Treatment effects are identified in a narrow neighborhood around each cutoff. As such, the estimates may not generalize to schools far from the thresholds or to districts with different administrative environments. Additionally, when covariates are added through the kernel, *multicollinearity with the running variable* and *increased input dimensionality* can inflate posterior uncertainty at the boundary, especially with sparse observations near c. Thus, attenuation from significant to non-significant may reflect either genuine **bias correction** (covariates explain the jump) or **power limitations** (wider intervals), and we refrain from strong directional claims in those cases.

8.2 Threats to Identification Assumptions

The validity of the regression discontinuity design (RDD) relies on two main assumptions: (1) continuity of potential outcomes in the absence of treatment at the cutoff, and (2) no precise manipulation of the running variable around the threshold. In this context, the FRPM and UPP eligibility percentages are derived from administrative data and largely reflect demographic and economic conditions, which limits the likelihood of strategic manipulation by schools. This is supported by McCrary density tests at the 35% and 75% FRPM thresholds, which reveal no significant discontinuities in the distribution of the running variable.

A more serious concern is measurement error, particularly due to rounding of FRPM percentages to the nearest whole number. These mass points can lead to clustering around integers, which may bias polynomial fits in classical RDD estimators or introduce unwanted structure in GP-RDD kernels. To address this, we use GP-RDD's nonparametric flexibility, which places a smoothness prior over the conditional outcome surface and can better accommodate such artifacts.

The validity of local randomization hinges on the absence of correlated confounders near the cutoff. Although covariate balance checks suggest that observable school characteristics do not jump discontinuously at the thresholds, unobservable factors such as staff discretion in program take up or reporting practices may still introduce noise. These limitations underscore the importance of robustness checks and triangulation with multiple estimation strategies.

A final limitation is that GP-RDD assumes constant, Gaussian errors [CKH2X]. In their simulations, GP-RDD performs well across many heteroskedastic scenarios, with coverage, interval length, and RMSE generally robust. The only exception arises in a specific pattern of declining variance toward the cutoff with large samples and strong heteroskedasticity, where intervals widen but coverage remains conservative and RMSE preferable.

8.3 Data Limitations

Several practical limitations in the data also constrain inference. First, BTB participation records are incomplete or inconsistently coded in some years, limiting the precision of treatment indicators. Second, public absenteeism files suppress values for schools with low student counts, particularly affecting small or alternative schools. Lastly, the FRPM variable, while mechanically calculated, may not capture the full extent of economic disadvantages.

CHAPTER 9

Conclusion and Future Work

9.1 Conclusion

This thesis has examined the causal effect of school level FRPM eligibility thresholds on chronic absenteeism in the Los Angeles Unified School District (LAUSD) using both classical and Bayesian statistical methods. Specifically, we employed Regression Discontinuity Design (RDD) at the 35% and 75% FRPM cutoffs, which we use as identification points for estimating the effects of various school-based services, including the Community Eligibility Provision (CEP) and the "Beyond the Bell" (BTB) program.

The primary methodological contributions include the combined use of classical local linear RDD with triangular kernels and robust bias correction, as well as a novel implementation of Gaussian Process RDD (GP-RDD), a Bayesian nonparametric extension that offers improved handling of uncertainty, heteroskedasticity, and discretized running variables.

Both methods yielded substantively similar treatment effect estimates, bolstering the credibility of the causal findings. The GP-RDD approach, in particular, allowed for endogenous bandwidth selection through marginal likelihood optimization and provided posterior credible intervals that more fully reflect modeling uncertainty. This Bayesian extension proved especially useful in accommodating the presence of integer-valued mass points in the FRPM variable and in allowing for smooth, flexible response functions without the need to specify polynomial order.

However, the results also reveal an unexpected and concerning dimension: at the 35%

FRPM threshold, eligibility is associated with declines in academic proficiency and increases in the share of students not meeting standards, particularly in mathematics. This challenges the assumption that crossing into eligibility automatically improves outcomes, and suggests that newly qualifying schools may require more immediate, intensive, or targeted interventions to translate additional resources into measurable academic gains.

9.2 Suggestions for Future Research

Several avenues for future work emerge from this analysis. First, the dataset could be extended to include additional academic years as they become available, allowing for dynamic causal inference using longer panel structures. This would make it possible to implement panel regression discontinuity and explore whether effects persist, diminish, or accumulate over time.

Second, the range of outcome variables could be broadened beyond chronic absenteeism. Future analyses might examine standardized test scores, graduation rates, disciplinary actions, or social emotional learning indicators. This would offer a more comprehensive view of how threshold based policies affect student well being and educational attainment.

Third, the FRPM thresholds examined in this study are not the only policy discontinuities available in education policy. Researchers could apply similar designs to other cutoff based programs, such as Title I funding, charter authorization caps, or per pupil incentive schemes to generalize the methodological approach across contexts or different data characteristics.

Fourth, there could be a delay in the treatment and the response to the treatment. For example, although a school may be eligible for funding in one year it does not necessarily fully translate to immediate funding availability. There are still many steps between funding eligibility, funding granted, funding disbursement, and funding making into the school such that it can affect the students. Future works could examine various delays in these treatment and responses. This study omitted this due to limited access of data on student outcomes in subsequent years after treatment. In the same vein, other statistical methods, perhaps other segmented regression techniques, could be used to detect cut-offs as a way to gain additional information on this delay.

Finally, other methods could be employed to capture geographic spillovers, such as the influence of neighboring schools' program access on a given school's performance.

Outside the educational context, the techniques developed here could be applied to other social policy domains that involve eligibility thresholds such as Medicaid income caps, Supplemental Nutrition Assistance Program (SNAP) guidelines, or housing voucher prioritization. As interest in evidence based policymaking grows, quasi-experimental methods like RDD and its Bayesian variants will continue to play a vital role in assessing the fairness and effectiveness of public programs.

REFERENCES

[AI22] Susan Athey and Guido W. Imbens. "Design-based Analysis in Difference-in-Differences Settings with Staggered Adoption." *Journal of Econometrics*, 2022. Emphasizes design-based identification strategies and importance of pre-analysis plans.

[AL99] Joshua D. Angrist and Victor Lavy. "Using Maimonides' Rule to Estimate the Effect of Class Size on Scholastic Achievement." *The Quarterly Journal of Economics*, 1999.

[Ble12] Myra Bleeker et al. "Study of School Breakfast After the Bell in Large Urban District." *Journal of School Health*, 2012. Urban district finding 0.10 SD gains in math and reading for low-income and Hispanic students.

[Can23] Seda Canbolat. "Attendance-Based Early Warning Systems: A Multi-Cutoff Regression–Discontinuity Application." *Education Economics*, 2023.

[CKH2X] Soonhong Cho, Doeun Kim, and Chad Hazlett. "Inference at the Data's Edge: Gaussian Processes for Estimation and Inference in the Face of Model Uncertainty, Poor Overlap, and Extrapolation." *Cambridge Medium*, 202X. Forthcoming.

[CM03] Janet Currie and Enrico Moretti. "Mother's Education and the Intergenerational Transmission of Human Capital: Evidence from College Openings." *The Quarterly Journal of Economics*, **118**(4):1495–1532, 2003.

[CS21] Brantly Callaway and Pedro H. C. Sant'Anna. "Difference-in-Differences with Multiple Time Periods." *Journal of Econometrics*, 2021. Introduces generalized DiD estimator accommodating treatment timing variation.

[CTN22] Andrés Cuadros-Meñaca, Michael R. Thomsen, and Rodolfo M. Nayga Jr. "The Effect of Breakfast After the Bell on Student Academic Achievement." *Economics of Education Review*, **86**, 2022. Arkansas study finding modest negative effect on math achievement and null effect on ELA.

[DWP10] Joseph A. Durlak, Roger P. Weissberg, and Molly Pachan. "A Meta-Analysis of After-School Programs That Seek to Promote Personal and Social Skills in Children and Adolescents." *American Journal of Community Psychology*, **45**:294–309, 2010.

[Dyn04] Mark Dynarski et al. "When Schools Stay Open Late: The National Evaluation of the 21st-Century Community Learning Centers Program." *U.S. Department of Education*, 2004.

[Edu21a] California Department of Education. “Local Control Funding Formula Frequently Asked Questions.”, 2021. Accessed July 2025.

[Edu21b] California Department of Education. “Title I, Part A School Allocations.”, 2021. Accessed July 2025.

[Goo21] Andrew Goodman-Bacon. “Difference-in-Differences with Variation in Treatment Timing.” *Journal of Econometrics*, 2021. Shows two-way fixed effects estimator decomposes into weighted averages, clarifies biases.

[GZ14] Craig Gundersen and James P. Ziliak. “Childhood Food Insecurity in the U.S.: Trends, Causes, and Policy Options.” *The Future of Children*, **24**(2):1–19, 2014.

[Inf21] California Legislative Information. “Education Code - EDC § 64001(a).”, 2021. Accessed July 2025.

[Jen23] Jeffrey M. Jenson et al. “Evaluating the Impact of Community-Based Afterschool Programs on Academic and Behavioral Outcomes.” *Children and Youth Services Review*, 2023. Matched comparison study showing positive impacts on reading and attendance.

[JL04] Brian A. Jacob and Lars Lefgren. “Remedial Education and Student Achievement: A Regression–Discontinuity Analysis.” *The Review of Economics and Statistics*, 2004.

[Kla08] Wilbert van der Klaauw. “Regression–Discontinuity Analysis: A Survey of Recent Developments in Economics.” *Journal of Econometrics*, 2008.

[LSW13] Jacob Leos-Urbel, Amy Ellen Schwartz, Margery Weinstein, and Sean Corcoran. “Not Just for Poor Kids: The Impact of Universal Free School Breakfast.” *Economics of Education Review*, 2013. NYC universal free breakfast increased participation but yielded null effects on attendance and achievement.

[Man12] Marco Manacorda. “Using Grade Retention to Estimate the Effect of Education on Dropouts in Uruguay.” *Journal of Labor Economics*, 2012.

[Rit24] Elizabeth Ritchie et al. “Systematic Review of Universal Free School Meal Programs.” *JAMA Network Open*, 2024. Synthesizes six studies; finds consistent increases in meal participation but small or null effects on attendance and academics.

[Soh22] John Sohn et al. “Performance-Based Funding and Student Achievement in South Korea: A Regression–Discontinuity Approach.” *Economics of Education Review*, 2022.

[SR20] Amy Ellen Schwartz and Micah Rothbart. "Let Them Eat Lunch: The Impact of Universal Free Meals on Student Performance." *Journal of Policy Analysis and Management*, 2020. Difference-in-differences analysis of universal free school lunch in NYC, finding test score improvements of 0.08 SD in math and 0.06 in ELA for middle-schoolers.

[SWW17] Guido Schwerdt, Martin R. West, and Marcus A. Winters. "The Effects of Florida's Third-Grade Retention Policy: Evidence from Panel Data." *American Educational Research Journal*, 2017.

[VRP07] Deborah Lowe Vandell, Elizabeth R. Reisner, and Kim M. Pierce. "Outcomes Linked to High-Quality Afterschool Programs: Longitudinal Findings from the Study of Promising Afterschool Programs." *Policy Studies Associates*, 2007.