Essays in Education and Crime

by

Emily Arnesen

A dissertation accepted and approved in partial fulfillment of the
requirements for the degree of
Doctor of Philosophy
in Economics

Dissertation Committee:
Jonathan MV Davis, Co-Chair
Michael Kuhn, Co-Chair
Benjamin Hansen, Core Member
David Liebowitz, Institutional Representative

University of Oregon
Spring 2025

DISSERTATION ABSTRACT

Emily Arnesen

Doctor of Philosophy in Economics

Title: Essays in Education and Crime

This dissertation contains three essays in applied microeconomics. Chapter I provides summaries of all three essays. Chapters II and III concern the economics of crime, while Chapter IV concerns the economics of education.

Chapter II, co-authored with Kyutaro Matzusawa, examines the effects of the legalization of sports betting on intimate partner violence. Using the exogeneity of NFL home-team upset losses interacted with the policy of sports betting legalization, we find that when an NFL team experiences an upset loss, intimate partner violence increases. We conduct heterogeneity analysis to show that the effect is larger where mobile betting is legal, in places where the total per capita bets are higher, during pay weeks and when a team is on a winning streak.

Chapter III, co-authored with Jonathan MV Davis and Tracey Meares, provides experimental evidence on the efficacy of youth outreach forums conducted in the Chicago Juvenile Detention Center (JTDC). We find that the program led to a reduction in new spells in the JTDC and an increased attachment to school.

Chapter IV explores how school districts allocate their expenditures following a school finance reform. I find that increased school district revenues and expenditures are associated with increased spending on positions which require a

teaching certification. I also find evidence that increases in school district revenues is associated with retaining teachers at a higher rate.

This dissertation includes both previously published, unpublished, and co-authored material.

CURRICULUM VITAE

NAME OF AUTHOR: Emily Arnesen

GRADUATE AND UNDERGRADUATE SCHOOLS ATTENDED:
University of Oregon, Eugene, OR
Gonzaga University, Spokane, WA

DEGREES AWARDED:
Doctor of Philosophy, Economics, 2025, University of Oregon
Master of Science, Economics, 2021, University of Oregon
Bachelor of Science, Economics, 2020, Gonzaga University

AREAS OF SPECIAL INTEREST:
Economics of Education and Crime

PROFESSIONAL EXPERIENCE:

Graduate Employee, Economics Department, University of Oregon, 2020-2025

GRANTS, AWARDS AND HONORS:

Graduate Student Paper Award, NAASE, 2024

Graduate Teaching Award, University of Oregon Department of Economics, 2024

Kleinsorge Summer Fellowship, University of Oregon Department of Economics, 2024

First Year Econometrics Award, University of Oregon Department of Economics, 2021

Dale Underwood Outstanding Graduate Student Scholarship, University of Oregon Department of Economics, 2021

Promising Scholar Award, University of Oregon, 2020-2021

PUBLICATIONS:

Davis, J. M. V., & Meares, T. & Arnesen, E. (2024). Improving Programming in Juvenile Detention: The Impact of Project Safe Neighborhoods Youth Outreach Forums.*Journal of Quantitative Criminology*, 41.1, 23-50.

ACKNOWLEDGEMENTS

I thank all of my dissertation committee members for helping me through my doctoral journey. I am indebted to Jonathan MV Davis for helping push my ideas forward and engaging me in research early on in the program. I am grateful for Michael Kuhn, who always asked me questions that made me think critically about my methodology and for guiding me through the process of earning a PhD. I am thankful for Benjamin Hansen, who has pushed me to present my research outside of the Department of Economics and has given advice on a lot of my work. I also want to express my gratitude to David Liebowitz for expressing interest in all my research, providing helpful comments, and taking time to serve on my committee.

To all the people who supported me throughout my journey over the last five years, thank you for your support. To my mom, who always listened when I called to talk about research or life. Your constant support helped sustain me through this process. To my dad, who passed away during this program, whose perseverance in encouraging my education and helping me make connections will never be forgotten. To my friends, thank you for understanding that school gets hectic, but have always welcomed me back with open arms when I have the time. Lastly, to my fiancé, Clay, your constant support and encouragement in everything I do and your kindness and compassion has shaped me to be the teacher scholar that I am today.

TABLE OF CONTENTS

Chapter Page

Chapter Page

LIST OF FIGURES

LIST OF TABLES

Table Page

Table | Page

CHAPTER I

INTRODUCTION

This dissertation contains four chapters. This chapter, Chapter I, provides an overview of the other three chapters. Chapters II and III are related to the economics of crime, while Chapter IV relates to the economics of education.

Chapter II, co-authored with Kyutaro Matzusawa, examines the relationship between legalized sports gambling, unexpected emotional cues, and reported intimate partner violence. A 2018 U.S. Supreme Court Ruling gave states the ability to legalize sports betting. Many states took advantage of this, and to date, 39 states and Washington D.C. have some form of legal sports betting. Using the unexpected emotional cue of an NFL team upset loss, we find that the sports betting increases intimate partner violence rates when an NFL team experiences an upset loss.

Chapter III contains previously published co-authored material with Jonathan MV Davis and Tracey Meares. This study is an program analysis of an RCT that took place in the Chicago Juvenile Detention Center. This analysis provides suggestions for how to conduct RCTs in juvenile detention centers when youth stays are relatively short. This analysis also finds that the youth outreach program in this study decreased recidivism rates into the Chicago Juvenile Detention Center for 8 months following the intervention and increased attachment to school.

Chapter IV analyzes school district expenditure distributions following a state-wide school finance reform. This study shows that school districts are predicted to receive more apportionment funds, they spend more on teacher salaries. This chapter also provides evidence that teachers in school districts that were estimated to receive more apportionment funds under the new funding formula stayed at higher rates.

CHAPTER II

SPORTS BETTING LEGALIZATION AMPLIFIES EMOTIONAL CUES & INTIMATE PARTNER VIOLENCE

This chapter is co-authored with Kyutaro Matzusawa. We both contributed to formulating the idea, data collection, analysis, coding, and writing.

2.1 Introduction

Family violence or intimate partner violence (IPV) is a huge problem both in the United States and worldwide, as approximately 41% of women in the United States and 26% of women worldwide have experienced some form of IPV in their lifetime (Centers for Disease Control and Prevention, 2024; National Domestic Violence Hotline, 2023; UN Women, 2023). Previous studies have demonstrated that experiencing family violence victimization can lead to adverse effects on mental health, resulting in lower cognitive ability scores, reduced employment opportunities, and diminished earnings (Adams-Prassl et al., 2023; Bhuller et al., 2022). Peterson et al. (2018) estimate that the lifetime cost of a female victim of IPV – from health, lost productivity, and criminal justice costs– is approximate $137,000 in 2024 dollars. Finally, research has documented that family violence victimization can generate negative externalities affecting other peers (Carrell and Hoekstra, 2010, 2012; Carrell et al., 2018).

One explanation for increased IPV is the role of emotional cues stemming from various causes and unexpected events. Card and Dahl (2011) leverage a potential emotional cue with pre-game NFL spreads and find that an upset loss (a team lost when they were predicted to win by more than 3.5 points) increases male-to-female IPV by up to 11.2 percentage points. Cardazzi et al. (2022) find similar patterns when examining the NBA, suggesting that Card and Dahl (2011)'s findings do not pertain to the NFL, where the nature of sports is more violent than other sports and the games are played only once a week. Arenas-Arroyo et al. (2021) and Beland and Brent (2018) find that negative emotional cues outside of sports, such as the COVID-19 pandemic or traffic congestion, also contribute to increased family violence. Conversely, Collins (2022) finds that positive emotional cues, such as voting for the winning presidential candidate, can reduce IPV.

Household finances can also influence family violence, although the direction of this relationship remains unclear. Several studies have found an increase in family violence during periods of negative economic shocks, such as the Great Recession (Schneider et al., 2016), COVID-19 pandemic (Arenas-Arroyo et al., 2021), mass layoffs (Lindo et al., 2018), and stock market losses (Lin and Pursiainen, 2023). These findings are consistent with the possibility of increases in the likelihood of conflicts arising from household finances (Carr and Packham, 2021). In contrast, a signaling model by Anderberg et al. (2016) and empirical analysis using data from the UK suggest that a higher risk of unemployment and lower wages decrease male IPV rates. Their findings suggest that factors like fear of job loss, the economic incentive to avoid divorce, and the associated loss of spousal insurance may play significant roles. This finding is also consistent with Aizer (2010)'s household bargaining model, which finds that decreases in the female-to-male wage gap reduce violence against women.

Participation in gambling can also impact the probability of committing an instance of IPV.[1] Gambling participation can result in unexpected outcomes where

[1]Several papers document a strong association between problem gambling and IPV. For example, Dowling et al. (2016) conduct a meta-analysis of six studies and find that 36.5% of

individuals will feel more negative emotional cues. Moreover, the gambling market will lead to lower expected but larger variations in household finances due to potential short-term gains and losses but the casinos making profits in expectation.

In 2018, a United States Supreme Court decision allowed states to legalize sports gambling for the first time since 1992.Following this ruling, 38 states and the District of Columbia legalized sports gambling. This legalization led to vastly increased interest and participation in sports gambling (American Gaming Association, 2023; Baker et al., 2024; Hollenbeck et al., 2024; Huble, 2023; Oxford Economics, 2017). Thus, in theory, sports gambling legalization can lead to changes in emotional cues and IPV. However, to our knowledge, no studies have investigated such causal relationships. In this paper, we are the first to fill this gap in the literature.

Using data from the 2011-2022 National Incident-Based Reporting System (NIBRS) and extending Card and Dahl (2011)'s model, we document that in states where sports betting has been legalized, the effect of upset losses on IPV is about 6 to 7 percentage points larger than in states without sports betting. Supplemental tests examining the effect on Sunday morning crimes, bar fights, or other non-IPV assaults confirm that changes in crime reporting or crime displacements do not drive this estimate. Furthermore, we find that the effects are driven by home teams on a winning streak, states with legal mobile betting, Sundays right after paydays, and states with a larger betting market. The pattern of these findings confirms that the reaction to gambling loss explains our results.[2] Together, our paper sheds some

problem gamblers are perpetrators of IPV. Korman et al. (2008a) find that 55.6% of problem gamblers had perpetrated IPV.

[2]Another potential mechanism noteworthy that may be at play is that sports gambling legalization can increase emotional attachments to the fans' favorite team. This heightened

light on one of the unintended consequences of legalized sports betting: Legalized sports betting can amplify negative emotional cues.

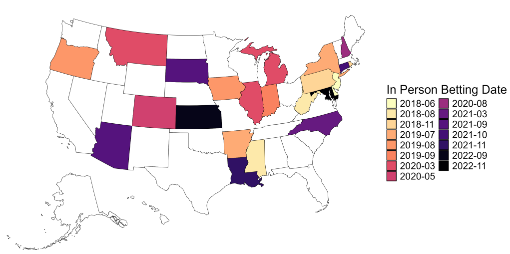
The remainder of the paper is organized as follows. Section 2.2 provides a brief history of sports betting in the United States. Section 2.3 explains each of the data sources used. Section 2.4 discusses the methodology used in our analysis. Section 3.5 provides results. Section 2.6 concludes.

2.2 Background

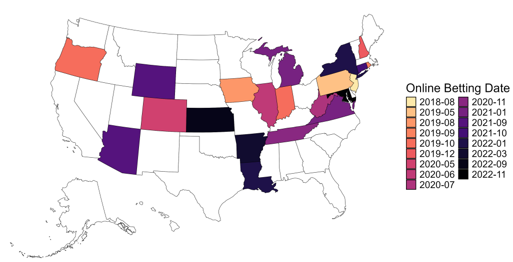
Sports gambling's availability was regulated by the Professional and Amateur Sports Protection Act of 1992 (PASPA). Before then, some states offered betting for big games such as the Super Bowl and considered legalizing sports betting in the 1980s. In 1992, the federal government stopped sports betting in all states except Nevada through the PASPA. For the next 26 years, Nevada was the only state where sports betting was legal. Delaware, Oregon, and Montana also offered some types of sports betting before PASPA, so what they had legalized before PASPA was grandfathered into the law. Delaware had legal parlay bets for football on three teams or more but stopped offering it in 1976. Oregon also offered parlay betting before PASPA but ceased offering it in 2007. Montana had legal sports pools, fantasy sports leagues, and sports tab games. However, the advent of the internet introduced new challenges. Online betting has become popular regardless of the state in which a person lives. The Unlawful Internet Gambling Enforcement Act of 2006 was passed, which did not ban online gambling, but banned the collection of credit, checks and electronic fund transfers to gambling websites in the United States.

attachment can lead to a larger adverse reaction to an upset loss of their favorite team. However, we do not directly test this hypothesis in this paper due to data limitations.

On May 14, 2018, in a court case *Murphy v. National Collegiate Athletic Association,* the U.S. Supreme Court overturned the PASPA. After this ruling, 38 states plus Washington D.C. made some form of sports gambling legal, and the sports betting market began growing with an estimated annual total revenue of $10 billion. There are two distinct ways in which people can bet on sports. Retail or in-person betting occurs within a physical building that a state's government has approved. These places are usually casinos; however, in some states, bars and stadiums have been approved to facilitate sports gambling. The other way people can bet on sports is through online (sometimes called mobile) betting. In most states where online betting is allowed, people can place bets from anywhere in the state as long as they have access to the internet.[3] A majority of states who have legalized some form of sports gambling legalized retail betting first. Fig. 1a and Fig. 1b show the month and year when each state launched retail and online sports gambling.



(a) Retail Gambling



(b) Online Betting

FIGURE 1. Sports Gambling Launch Date
Notes. These figures show the month and year when each state launched legalized retail or online (mobile) sports gambling between 2018 and 2022. We exclude NM, ND, WA, and WI, which legalized sports gambling only on tribal land.

[3] There are some states, such as Washington, which only allow online sports gambling on tribal land.

2.3 Data

Crime Data

We obtain our crime data from the National Incident-Based Reporting System (NIBRS). The NIBRS collects incident-level crime data reported by around 8,500 local law enforcement agencies, covering approximately 146.5 million people in the U.S.[4] The NIBRS data has several attractive features that we can utilize. First, we can observe the number of incidents rather than arrests.[5] Because the probability of arrest may be a function of police discretion, using arrest may underestimate the actual crime rate. Second, it permits the identification of the date and time of offense, which we can leverage to estimate the hourly crime rates during the game day. Finally, the rich set of information provided in the NIBRS data allows us to identify whether a crime was an IPV.

Following Card and Dahl (2011), our primary variable of interest is defined as any male-to-female simple assault, aggravated assault, or intimidation that occurred against a spouse or partner at home during a 12-hour time window (12 pm to 11:59 pm Eastern Time) centered around game and post-game window.[6] Because the rate of violence varies substantially across the days of the week and most NFL games occur during Sundays, we restrict our sample to the 17 Sundays during the regular NFL season.[7] To match the length of Card and Dahl (2011)'s 12-year sample window, we focus our sample on the years 2011-2022.

[4]Our NIBRS data comes from Kaplan (2023) and can be downloaded from ICPSR.

[5]Approximately only half of reported incidents result in an arrest.

[6]Approximately half of the NFL have a start time of 1 pm Eastern Time.

[7]Because NIBRS data is not available for 2023, for the 2022 season, we restrict our sample to 16 Sundays.

One limitation of the NIBRS data is that agencies may not consistently report across the year. For instance, an agency may report all the incidents in January but not February. Thus, interpreting zero crime can pose a challenge because we cannot distinguish whether the crime rate is truly zero or not reported in the data. To address this concern, we follow Card and Dahl (2011) and restrict our agencies where they report any crime during 13 out of 17 game-days during a season.[8] By restricting the sample to agencies that report any crime during most game-days, we can be more confident that zero family violence reports reflect an actual absence of crime rather than a reporting artifact.

Another limitation of the NIBRS is that agencies will enter the NIBRS sample and, in some rare cases, stop reporting at different points in time. Hence, the changes we observe in county-level crime rate may be driven by changes in the agencies reporting to the NIBRS. To circumvent this issue, we follow Lindo et al. (2022) and aggregate our sample to the agency level rather than a more aggregate level such as the county level.

Game Spread Data

NFLOddsHistory.com provides a final pre-game point spread for each NFL game, which measures the likelihood that a team will win a game using an algorithm.[9] A negative spread favors the team to win, and a larger absolute value implies a higher likelihood of winning or losing. For example, a point spread of -9 implies that the team must win by more than 9 points for a bet on that

[8]We also follow Card and Dahl (2011) and exclude college or special agencies. Our estimates are robust to different samples, such as including these agencies and using different numbers of days reported.

[9]Data can be obtained from https://www.sportsoddshistory.com/nfl-game-odds/

team to pay off, and that the team is favored to win by a quite large margin. Previous research has shown that the point spread is an unbiased predictor of game outcomes (Gandar et al., 1988; Pankoff, 1968).

We match the pre-game point spread with the actual outcome of the game to create a measure of unexpected wins and losses. Following Card and Dahl (2011), we define an unexpected loss (upset loss) if a team had a point spread of -4 or more (favored to win), but lost the game; unexpected win (upset win) if a team had a point spread of 4 or more (favored to lose), but won the game; and close loss if a team had a point spread between -4 and 4 (close game) and lost the game. In Table A.1, we present the total count and percentage of unexpected outcomes for each NFL teams. On average, each team experiences an unexpected loss about once per season (approximately 27% of the predicted to win games), close loss about 3 times per season (approximately half of the predicted close games), and upset win about once per season (approximately 28% of the predicted to lose games).

Mapping Home Team

Mapping each NIBRS agency to an NFL franchise can be challenging because not all large cities with an NFL stadium have data for NIBRS. This sample limitation can result in loss of power or generalizability of our treatment effects. To address this issue, Card and Dahl (2011) focus only on states with a single NFL team (or nearby team) with at least four years of crime data, and assign all jurisdictions within a state to a team in the same state.[10] Panel a of Fig. A.1 shows the states with at least one NIBRS agency with an NFL team. However, a

[10]For nearby teams, they assign South Carolina to the Carolina Panthers and New Hampshire and Vermont to the New England Patriots.

limitation of this approach is that the estimated treatment effects only apply to jurisdictions with exactly one NFL team.

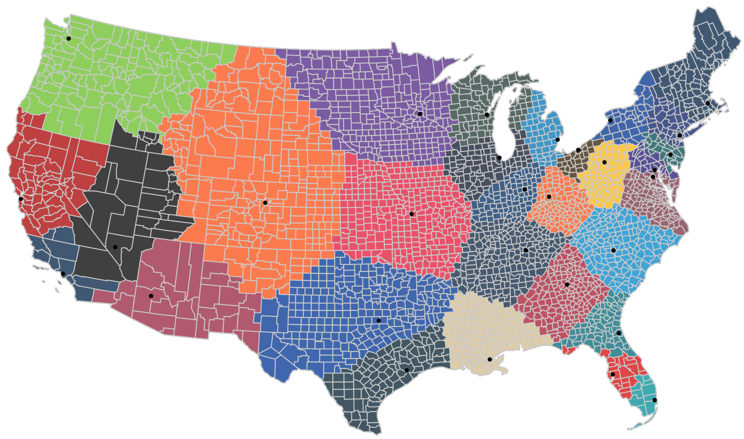
To extend upon Card and Dahl (2011)'s model, our preferred approach of defining a "home" NFL team is to map each county with at least four years of crime data based on NFL fandom. Our preferred method of defining NFL fandom is based on proximity. Specifically, we find the closest NFL stadium in linear distance from its centroid for each county.[11] Panel a of Fig. 2 presents the "home team" for every county in the U.S. Using this approach, our total selection includes 2447 police agencies within 831 counties across 40 states (see Panel b).

Fig. 3 plots the mean IPV rate per 100,000 when mapping the home team using our preferred method. Panel (a) suggests that relative to when a home team wins as expected, the mean IPV rate is 5% higher (0.412 vs. 0.439) when the team experiences an upset loss. However, when comparing close wins vs. close losses or losses as expected vs. upset wins, the difference in the mean is not stark. Strikingly, panel (b) shows that these differences in mean IPV rates across upset loss days vs. non-upset loss days are higher when sports betting is legal (19% vs. 3.6% increase).

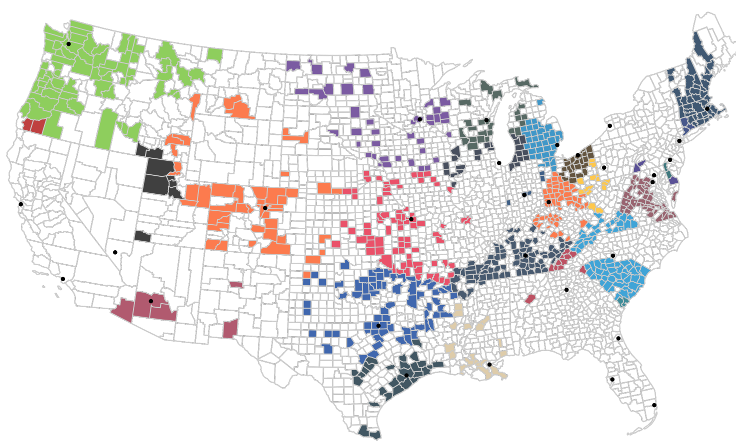
[11] Two stadiums have two home teams: Los Angeles Rams and Chargers, and New York Jets and Giants. Since California is not in our sample, we ignore the former stadium. We use the New York Giants for the New York teams since they are a more popular franchise with a larger fanbase. However, our results are robust to (1) mapping counties to Jets instead of Giants or (2) excluding Giants from our analysis.

FIGURE 2. Home Team for Each County Based on Proximity

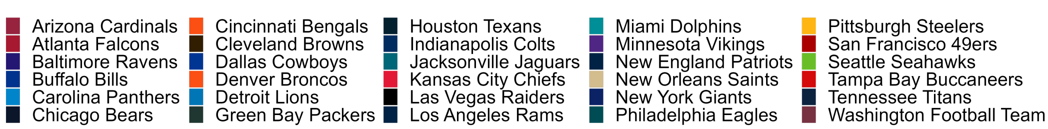
(a) All U.S. Counties



(b) All Counties Covered Under NIBRS



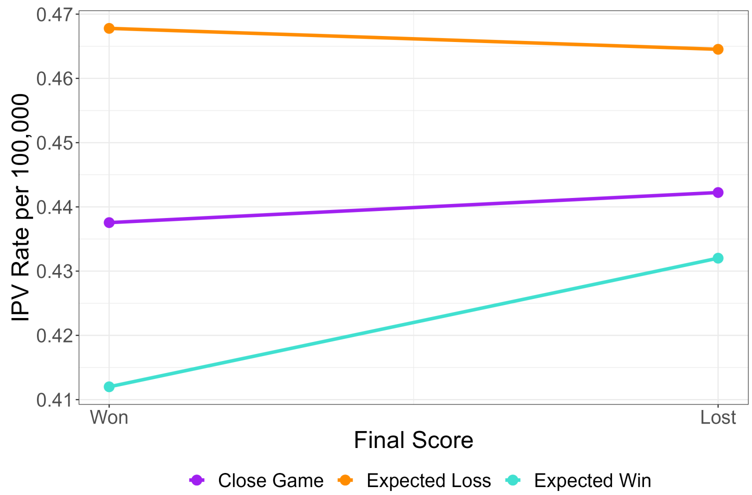
(c) Legend



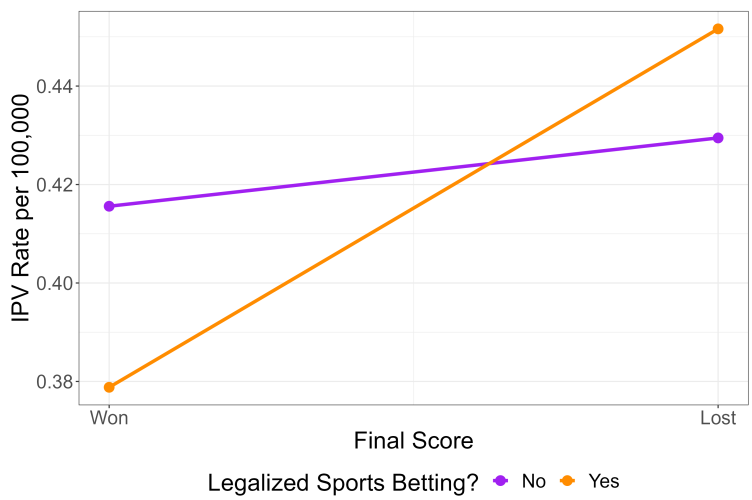
Notes. These figures show the closest NFL team based on linear distance for each continental U.S. county, based on stadium locations as of 2021. Panel (a) presents the home team for all counties; panel (b) shows the same for counties in the NIBRS sample. Black dots represent stadium locations.

FIGURE 3. Weighted Mean IPV

(a) By Expected Game Outcome & Actual Game Outcome



(b) By Sports Betting Legalization & Actual Game Outcome



Notes: These figures show population-weighted IPV rates per 100,000. Panel (a) plots the weighted mean by expected game outcome and actual game outcome. Panel (b) restricts the sample to teams that were expected to win and plots the mean by actual game outcome & sports gambling legalization.

Because proximity may not fully capture the intensity of the fanbase, we also experiment with supplemental definitions of the fanbase. Specifically, we obtain information for the most favored team for each county using data from

Seatgeek and Facebook. These measures are based on the total number of ticket searches and Facebook likes.[12] We also combine both these datasets to obtain the temporal changes in fandom and to estimate the intensity of fans based on different definitions of fans.[13] The advantage of such a method is that we can examine a more precise fanbase. Fig. A.1 Panels (b) and (c) show the fandoms for each county with at least one NIBRS agency. A caveat with combining both sources for fanbase is that because these datasets measure fanbase differently, our estimated effect can be driven by changes in fanbase because of changes in the actual measurement. For this concern, we also experiment with restricting the counties where the fanbase is consistent across both data sources.

2.4 Empirical Strategy

Our main empirical strategy builds off Card and Dahl (2011), but the main difference is that we also include interaction between unexpected game outcomes and sports betting legalization. More specifically, we estimate the following model:

[12] The data for SeatGeek is for the year 2018, and the data for Facebook is for the year 2013.

[13] This may be important because three teams (Rams, Chargers, and Raiders) moved cities during our sample period, which may change who they root.

$$
\begin{aligned}
IPV_{isw} = & \beta_1 \text{Exp Win}_{isw} + \beta_2 \text{Exp Loss}_{isw} + \beta_3 \text{Exp Close}_{isw} \\
& + \beta_4 \text{Upset Loss}_{isw} + \beta_5 \text{Upset Win}_{isw} + \beta_6 \text{Close Loss}_{isw} + \alpha_0 \text{Policy}_{isw} \\
& + \alpha_1 \text{Exp Win}_{isw} * \text{Policy}_{isw} + \alpha_2 \text{Exp Loss}_{isw} * \text{Policy}_{isw} \\
& + \alpha_3 \text{Exp Close}_{isw} * \text{Policy}_{isw} \\
& + \alpha_4 \text{Upset Loss}_{isw} * \text{Policy}_{isw} + \alpha_5 \text{Upset Win}_{isw} * \text{Policy}_{isw} \\
& + \alpha_6 \text{Close Loss}_{isw} * \text{Policy}_{isw} \\
& + \delta_i + \gamma_s + \phi_w + \rho \text{Holiday}_{sw} + \rho X_{isw} + \varepsilon_{isw}.
\end{aligned}
\tag{2.1}
$$

In Eq. (2.1), IPV_{isw} is the count of intimate partner violence (IPV) in agency (i), season (s), and NFL week (w). Exp Win_{isw} is a dummy variable indicating whether a "home" team in agency i is expected to win; Exp Loss_{isw} is a dummy variable indicating whether a "home" team is expected to lose; and Exp Close_{isw} is a dummy variable indicating whether a "home" team is expected to have a close game. Upset Loss_{isw}, Upset Win_{isw}, and Close Loss_{isw} are dichotomous variables indicating whether a game resulted in upset loss, upset win, or close loss, respectively. We also include agency (δ_i), season (γ_s), and week of the season (ϕ_w) fixed effects. We also include controls for whether a Sunday is a holiday or not.[14] X_{isw} is a vector of time-varying weather controls (daily average temperature and total precipitation). Finally, Policy_{isw} is a dichotomous policy treatment variable indicating if a state had launched any in-person or online sports betting.[15]

[14] Holiday includes Labor Day weekend, Columbus Day weekend, Halloween, Veteran's Day weekend, Thanksgiving day weekend, Christmas Eve and Day, and New Year Eve and Day.

[15] If Policy_{isw} is a constant, then Eq. (2.1) will be identical to the model estimated by Card and Dahl (2011).

Following Card and Dahl (2011), we restrict our counterfactual to any Sundays where there are no "home" games.[16] To account for zero counts, we estimate Eq. (2.1) using a poisson with agency population as the exposure variable. Finally, because our policy treatment variable varies across states and our game outcome varies by team and season, we cluster our standard errors for Eq. (2.1) at the state-by-game-by-season level.

One underlying assumption for Eq. (2.1) to produce an unbiased estimate is strict exogeneity. Formally, we require that the unexpected outcome we observe is as good as random, conditional on our observables. Under strict exogeneity assumption, we can still recover the causal α_4, α_5, and α_6 in Eq. (2.1) even if sports betting legalization is endogenous to the error term (Nizalova and Murtazashvili, 2016).[17] This assumption is likely to be valid because after controlling for the pre-game spread, the outcome of the game is random (Card and Dahl, 2011; Pankoff, 1968; Gandar et al., 1988).

Another assumption necessary for Eq. (2.1) is the stable unit treatment value assumption (SUTVA), which requires no spatial spillovers to states that did not legalize sports betting. Justifying this assumption is challenging because individuals can travel to nearby jurisdictions to gamble on sports or use VPNs for mobile betting. However, in the event of a violation of SUTVA, our estimate will be attenuated towards zero, and our estimated effect will serve as a lower bound and not a spurious relationship.

[16]A team can have no home games if they have a bye-week or play a game on a different day of the week (i.e., Thursday or Monday night). We document that approximately 20% of our sample has no game on a given Sunday.

[17]Because our interests are in the changes in behavior when a fan experiences an unexpected shock rather than a baseline shift as result of implementing a policy, we will not be arguing whether sports gambling policy is exogenous to our outcome of interest.

One potential threat to our identification is that our treatment turns on and off (i.e., unexpected game outcomes change heterogeneously across jurisdictions). This feature of our treatment can lead to each treated observation receiving unequal (and sometimes negative) weights. Consequently, it can lead to biased (and even different in sign) estimates than the true average treatment effects Chaisemartin and d'Haultfoeuille (2020). We note that this concern is minimal for two reasons. First, when we calculate the negative weights in our main specification, we find no observations receiving negative weights, suggesting that the sign of our estimated effect is the same as the sign of the true treatment effect. Second, when we experiment with Chaisemartin and d'Haultfoeuille (2020) estimator, which allows the treatment to turn on and off, we find contemporaneous effects of 0.071 per 100,000 (or 16.1 percent relative to the baseline mean of 0.44 with SE=0.039 and p-value=0.070) for our upset loss times betting coefficient, which is qualitatively similar to our estimated treatment effects.[18]

Our last assumption is the first-stage assumption. We assume that the intensity of sports gambling participation increased following the legalization of sports gambling. This assumption can be violated, for instance, if the legalization of sports gambling just moved individuals from the black market to the legal market. While the total count of individuals participating in the legal and illegal sports gambling cannot be directly measured, we believe this assumption is justified for three reasons. First, several reports and papers document evidence suggesting that the intensity of sports gambling, both in the extensive and intensive margins, has increased. For example, two recent working papers by Baker et al. (2024)

[18]Because these estimations and diagnostics do not allow for Poisson estimation, we conduct these using OLS where the outcome is IPV per 100,000. Our main model using OLS shows qualitatively similar results.

and Hollenbeck et al. (2024) use household-level transaction data and consumer credit data, respectively, and find causal evidence that sports betting expenditure increased and overall consumers' financial health deteriorated after the legalization of sports gambling. Moreover, van der Maas et al. (2022) find that the legalization of online sports gambling led to an increase in the activity in online gambling message board. Another paper by Humphreys (2021) finds that sports betting legalization had a cannibalism effect where the amount of money wagered in video game lottery decreased by $900 million in West Virginia, suggesting a potential change in gambling behavior as a result of sports gambling legalization. Finally, reports from American Gaming Association (2023) and Huble (2023) suggest that the number of adults in the U.S. who are interested in sports gambling increased by 24 million between 2019 and 2023 and that calls into the National Problem Gambling Helpline increased by 45% between 2021 and 2022.[19]

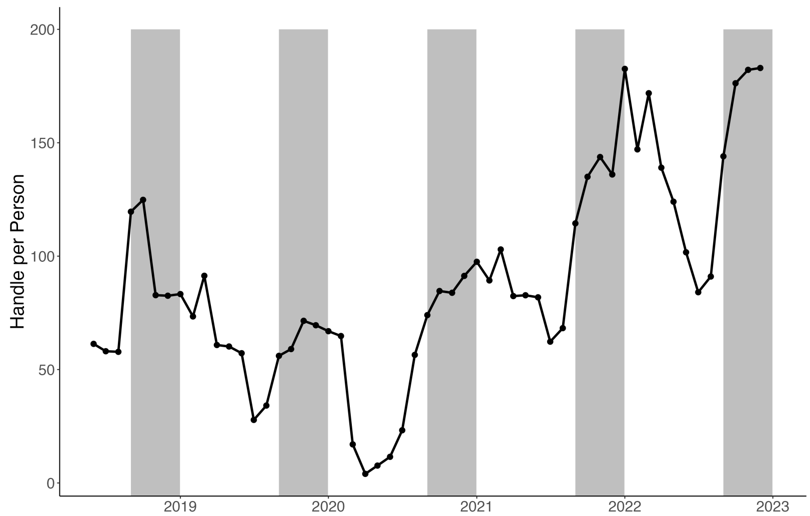
Next, we validate our first-stage assumption in Panel (a) of Fig. 4 by showing that the monthly money wagered on sports by average individuals over the age of 21 has been rising following the 2018 U.S. Supreme Court Case.[20] This figure confirms that legalized sports gambling indeed increased the size of the sports betting market, especially following the COVID-19 pandemic recovery, as the amount wagered increased by approximately 68 percent from $1,026 per person (in 2019) to $1,729 per person (in 2022). Moreover, we also show that participation in sports gambling spikes during football season (September to January). This pattern

[19] In the intensive margin, an online consumer survey data by EY noted that illegal sports bettors responded that the average amount wagered significantly increases when sports gambling is legal (Oxford Economics, 2017).

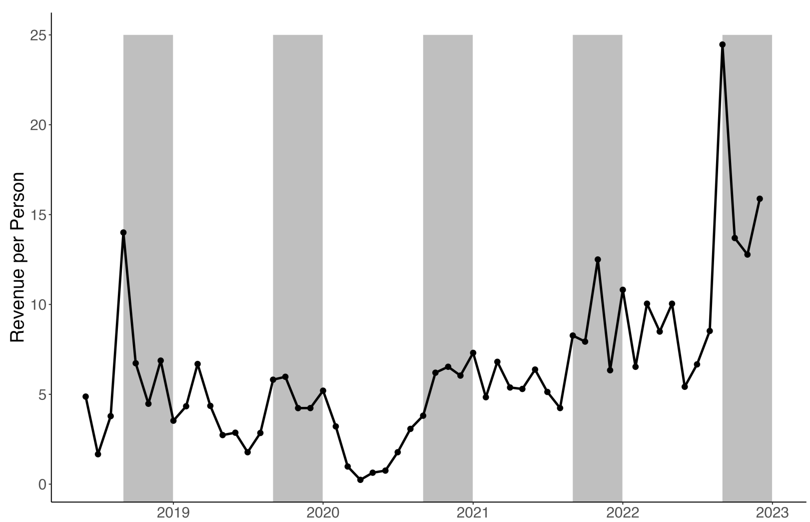
[20] In Panel (b), we show that the trend pattern is similar when we examine monthly revenue from sports gambling.

FIGURE 4. Total Handle & Revenue per Adult from Sports Betting

(a) Monthly Handle



(b) Monthly Revenue



Notes: These figures show monthly U.S. handle (panel a) and revenue (panel b) from sports betting. The data is derived from Legal Sports Report. The y-axis is scaled by dividing the total amount by the adult (21+) population in states where sports gambling was legal. The gray shaded area in panel (b) represents the NFL season.

is consistent with the fact that the NFL is one of the most popular sports to bet on and that the NFL is an appropriate setting for our study.[21]

Finally, we validate our assumption by examining if interest in sports gambling, as measured by data from Google Trends, increased after sports betting legalization. More specifically, we estimate the following model:

$$Searches_{it} = \sum_{j=-23}^{24} \beta_j D_{i,t}^{j} + \delta_i + \gamma_t + \varepsilon_{it} \tag{2.2}$$

In Eq. (2.2), $Searches_{it}$ is the total Google search for the keywords "sportsbook" or "sports betting" in state i for each month-year t.[22] We denote $D_{i,t}^{j}$ as our treatment leads and lags indicator, representing j months before or after sports betting became legal. δ_i and γ_t are our state and month-by-year fixed effects. We weight our estimates using state-level population and cluster our standard errors at the state level (Bertrand et al., 2004). We also experiment with using Sun and Abraham (2021) estimator using states that did not legalize sports betting by December 2022 as our counterfactuals to guard against negative weights arising from staggered treatment and dynamic treatment effects (Chaisemartin and d'Haultfoeuille, 2020; Goodman-Bacon, 2021; Sun and Abraham, 2021).

In the top row of Fig. 5, we present our TWFE event study estimates of Google Trends search for the keywords "sports betting" (panel a) and "sportsbook" (panel b). Though there may be some anticipatory effect in the pre-treatment period, which may be correlated with the announcement of the sports betting

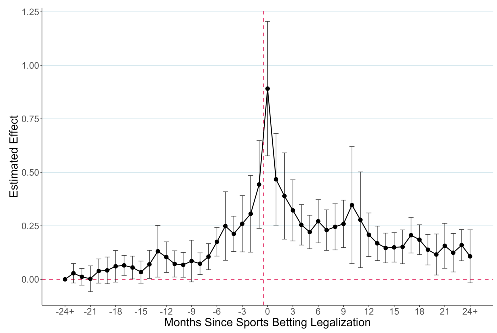
---

[21] We acknowledge that for more than half of the NFL season, the NBA is also in season, which will also increase betting revenue. While we cannot distinguish how much money is wagered per sport, it is estimated that the NFL is the most bet on sport, with 81% of sports bettors placing an NFL wager (Walsh, 2023).

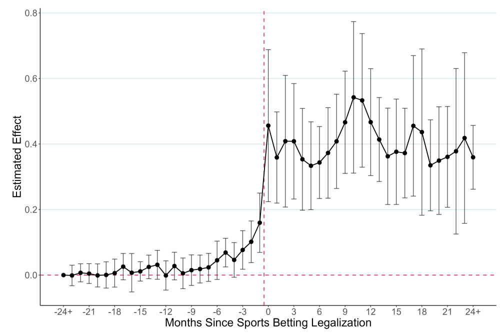
[22] Sportsbook is a physical or online platform where one can place bets.

FIGURE 5. TWFE and Sun & Abraham Event Study Estimates of Google Trends Search

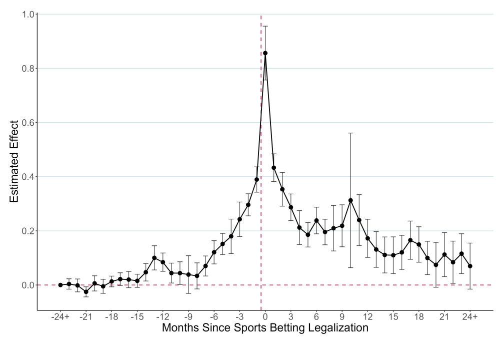
(a) “Sports Betting” – TWFE



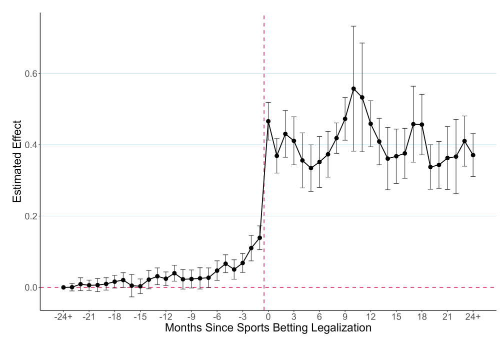
(b) “Sportsbook” – TWFE



(c) “Sports Betting” – Sun & Abraham



(d) “Sportsbook” – Sun & Abraham



Notes: These figures show two-way fixed effects (top row) and Sun & Abraham (bottom row) event study estimates of the impact of sports gambling legalization on Google searches from January 1, 2015 to December 31, 2022, for the keywords “sports betting” (left column) and “sportsbook” (right column). All estimates include state and month-by-year fixed effects, and are weighted by state-level population. Error bars show 95% confidence intervals based on standard errors clustered at the state level.

legalization, there is an apparent post-treatment reaction to the policy enactment. We find that around the month of the policy enactment, people are significantly more likely to search for the keyword “sports betting”. Furthermore, the search for the term “sportsbook” rose and persisted immediately following the policy enactment. The magnitude of the effect for the latter keyword suggests an approximately 0.37 percentage point increase (or 197 percent relative to the baseline mean of 0.189) in the post-treatment period. In the bottom row, we continue to find similar patterns of results when we use a different difference-in-differences technique. In summary, these estimates establish the first stage effect of sports betting legalization on interest in sports gambling.

2.5 Results

Sports Gambling, Emotional Cues, & IPV

In Table 1, we present our estimates when we interact our indicators for various game outcomes with an indicator for whether a state legalized sports gambling. We first document that our estimates are robust to the inclusion of various control variables and whether we use population as our exposure variable (i.e., weighted vs. unweighted). Our estimates imply that legalizing sports gambling increases the effect of upset loss on IPV by 4.1 to 6.31 percentage points ($\alpha_0 + \alpha_4$). The magnitude of our point estimates suggests that when sports betting is legal, IPV instances increase by 8.5 to 9.6% (or by 0.035 to 0.040 per 100,000 relative to the mean).[23] We also document a statistically significant change in individual's response to a home team's close loss when sports gambling is allowed. These findings support the possibility that an unexpected loss may lead to additional emotional cues and fans may express their frustration and anger towards other people (Hing et al., 2022; Korman et al., 2008b; Suomi et al., 2013).

[23] To derive the percent change, we add our upset loss coefficient, betting coefficient, and upset loss times betting coefficient, then we take the exponential and subtract one from it. For instance, column (1) implies exp(0.0263+0.0996-0.0406)-1=0.0890. To calculate level change, we multiplied our percent change by our baseline mean IPV rates of 0.417 per 100,000.

TABLE 1. Heterogeneous Treatment Effects of Legalized Sports Betting

	(1)	(2)	(3)	(4)	(5)
Upset Loss	0.0263	0.0298	0.0289	0.0303	0.0342*
	(0.0202)	(0.0198)	(0.0199)	(0.0198)	(0.0197)
Upset Loss*Betting	0.0996*	0.0969*	0.0993*	0.1000*	0.0975*
	(0.0509)	(0.0506)	(0.0508)	(0.0511)	(0.0513)
Close Loss	-0.0158	-0.0159	-0.0151	-0.0135	-0.0140
	(0.0128)	(0.0127)	(0.0126)	(0.0125)	(0.0124)
Close Loss*Betting	0.0783**	0.0825**	0.0811**	0.0764*	0.0772**
	(0.0384)	(0.0393)	(0.0393)	(0.0394)	(0.0387)
Upset Win	0.0236	0.0260	0.0256	0.0264	0.0259
	(0.0203)	(0.0201)	(0.0201)	(0.0199)	(0.0197)
Upset Win*Betting	0.0147	0.0126	0.0120	0.0095	0.0157
	(0.0641)	(0.0673)	(0.0670)	(0.0668)	(0.0653)
Betting	-0.0406	-0.0350	-0.0362	-0.0390	-0.0504
	(0.0350)	(0.0348)	(0.0347)	(0.0349)	(0.0341)
N	301,854	301,854	301,854	301,854	301,854
Agency FE?	Yes	Yes	Yes	Yes	Yes
Season FE?	Yes	Yes	Yes	Yes	Yes
Week FE?	No	Yes	Yes	Yes	Yes
Holiday FE?	No	No	Yes	Yes	Yes
Weather Control?	No	No	No	Yes	Yes
Unweighted?	No	No	No	No	Yes

Notes: This table reports interaction effects between expected game outcomes and sports betting legalization. All models include agency and seasonal fixed effects. Standard errors (in parentheses) are clustered at the agency level. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

While this estimated effect is large (over 100 percent increase relative to when states did not legalize sports betting), this magnitude is plausible for several reasons. First, when we examine the estimated effect of upset loss on only the states that legalized sports gambling, we find a beta coefficient of 9.1 percentage points, similar to that reported by Card and Dahl (2011). Second, there are several anecdotal examples of fans being upset about Fantasy Football, another form of sports gambling, or losses from their favorite team, and taking out their frustrations directly against players.[24] Third, the fact that problem gambling has adverse effect on IPV (i.e., 36.5 to 55.6 percent of gamblers who commit instances of IPV (Dowling et al., 2016; Korman et al., 2008a) and the fact that sports gambling legalization has significant adverse effects on problematic gambling behavior (i.e., 45 percent increase in calls to gambling hotline (Huble, 2023) imply that the risk of increased IPV may be significant. Fourth, a conservative estimate suggests about 22 million dollars wagered per NFL game (Edelstein, 2021).[25] If we assume that, on average, the bets are placed evenly across both teams, then this estimate suggests that one game, in expectation, leads to approximately 11 million dollars lost.

In Table A.2, we present the sensitivity of our findings to different definitions of "home team" and the inclusion of different geographic regions. In column 1, we follow Card and Dahl (2011)'s mapping strategy, focusing on the entire state with one NFL team. In columns 2 to 5 of Table 2, we experiment with different sources to define the fanbase and map home teams to the county level.[26;27] Our results for upset losses remain qualitatively similar when we use different geographies or

[24]Two recent examples from the 2023 NFL season include Alexander Maddison and Tyler Bass. When Maddison was underperforming and fans were losing their Fantasy Football game, some fans messaged Maddison on Instagram, calling racial slurs. In the latter example, Bass missed a critical play, leading to his team missing the playoffs. Consequently, he had to deactivate his social media because he received many online harassment and death threats.

[25]They estimate $350 million per week, which we divided by 16 games to estimate our per game number.

[26]Due to the relocation of the Rams, Raiders, and Chargers during our sample period, we are unable to identify the fanbase for these teams either before or after the move, so we exclude them from our analysis.

[27]For columns 5 and 6, we use Facebook likes to define our home team for 2010-2015 (pre-Rams move) and Seatgeek ticket searches for 2016 onwards.

definitions of a home team. We note that the coefficient on close loss interacted by sports betting loses precision, though the magnitude is still large and meaningful (3.5 to 6.8 percentage points).

To further assess the robustness of our estimates and address concerns about data cleaning and empirical choices, we present the results of sensitivity analyses in Table A.3. In column 1, we present our main coefficient for comparison. In columns 2 and 3, we demonstrate that our estimates remain robust when we require each agency to report data for multiple years. These findings imply that the changes in sample composition over time are not driving our results. Column 4 shows that our estimated effect is robust even when we exclude 2020, a period marked by the COVID-19 pandemic and potential issues with crime reporting. In column 5, we restrict our comparison to the treatment state and analyze the differences between the pre- and post-betting periods. The consistent patterns of results confirm that heterogenous treatment effects across states that legalized sports gambling vs. those that did not are not driving our results. Moreover, the result from column 5 implies that spatial spillovers and SUTVA violations are not much of a concern. Finally, in columns 6 and 7, we employ alternative specifications: negative binomial (column 6) and OLS estimates using the rate of IPV instances per 100,000 (column 7). Despite these changes, we find qualitatively similar results, suggesting that our findings are robust to different modeling approaches. In addition to Table A.3, we also experiment with excluding each state and team in Fig. A.2. The results from this exercise confirm that a particular state or team does not drive our estimated effect. Collectively, these sensitivity analyses provide reassurance that our estimated effect is not an artifact of specific data cleaning procedures or empirical choices but rather a robust finding in our study.

In Table 2, we extend our analysis to examine the effect of upset losses at various times of the day and on different types of assaults. Columns 1 to 4 focus on the timing of intimate partner violence (IPV) incidents: Sunday morning (before the game), Sunday afternoon (around or immediately after most games), Sunday night (around or immediately after the Sunday night game), and Monday mornings (the morning after). The findings in column (1) indicate no evidence of a change in IPV instances during the "pre-treatment" period, suggesting that there may be no spurious relationship between treatment (upset loss) and the rate of IPV. These findings provide some evidence that our treatment may be orthogonal to IPV outcomes. Columns 2 and 3 show larger effects on IPV occurring between noon and 5:59 pm compared to those occurring between 6 pm and midnight, although the 95% confidence intervals of these estimates overlap. In column 4, we find no evidence of IPV instances increasing on Monday mornings following an upset loss by a home team. These findings support the hypothesis that reactions to emotional cues and potential financial losses may be temporal and immediate rather than persisting over an extended period.[28]

In columns 5 and 6 of Table 2, we estimate the effect on other types of assaults: bar fights and non-IPV assaults. This exercise aims to detect any crime displacement resulting from NFL games. For example, if sports gambling creates an incapacitation effect, causing more people to stay home and watch the game instead of going out to bars, we might expect to see a negative effect on bar fights and other assaults. The point estimates show no negative decline in these types of assaults. Instead, we find a positive effect on upset loss, which is consistent with the idea that unexpected emotional cues can lead to more violent behavior beyond

[28]When we conduct our exercise focusing on just the morning games, we continue to find similar patterns of findings.

just IPV, especially in locations where alcohol people consume alcohol and watch sports. The findings in columns 5 and 6 suggest that displacement of the location where violence occurs is not a predominant channel.

TABLE 2. Estimated Effect on Other Time, Day, & Assaults

	(1)	(2)	(3)	(4)	(5)	(6)
Upset Loss	-0.0218	0.0578**	0.0115	0.0062	0.0693	0.0111
	(0.0233)	(0.0272)	(0.0281)	(0.0288)	(0.0923)	(0.0121)
Upset Loss*Betting	0.0744	0.1198	0.0807	-0.0537	0.4122*	-0.0010
	(0.0612)	(0.0767)	(0.0746)	(0.1006)	(0.2448)	(0.0248)
N	300,543	285,836	294,540	262,132	135,010	310,180
Outcome	IPV	IPV	IPV	IPV	Bar Fights	Non IPV Assaults
Day	Sun	Sun	Sun	Mon	Sun	Sun
Time	0to12	12to18	18to23	0to12	12to23	12to23

Notes: Standard errors are clustered at the state-by-season-by-team level. Bar fights are defined as any assaults occurring between 12 to 23 pm on Sundays at a bar or nightclub. Non IPV assaults are defined as any assaults occurring between 12 to 23pm on Sundays that is not classified as IPV. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Heterogeneity

Two potential mechanisms may explain why we observe the adverse effect of legalized sports gambling on the relationship between upset loss and IPV. The first channel is the financial loss from betting on one's favorite team, who is favored to win but ends up losing. The second channel is increased emotional attachment toward one's favorite team from participating in sports gambling. While the latter mechanism is more complex to tease out due to not having data on the intensity of the fanbase, we conduct several heterogeneity estimates to confirm that the former is a predominant channel that is driving our results.

In Table 3, we present our treatment effects after restricting our sample based on the home team's previous performance and by week of the month. To ensure consistency in our comparison group within our main specification, we restrict our primary sample to game days when a home team was expected to win.[29] As a result, our estimated effect can be interpreted as the effect relative to game days when a team won as expected.

TABLE 3. Heterogeneity Estimates by Home Team Winning Record and Week Day

		Won?			Covered Spread?			Payday?	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Upset Loss	0.0297	-0.0094	0.0499*	0.0546	0.0274	0.0233	0.0184	0.0273	0.0535*
	(0.0205)	(0.0337)	(0.0264)	(0.0338)	(0.0310)	(0.0289)	(0.0404)	(0.0327)	(0.0304)
Upset Loss*Betting	0.1261**	0.1283	0.1737**	0.2085**	0.0946	0.2085**	0.2464**	0.2428***	-0.0533
	(0.0506)	(0.0991)	(0.0707)	(0.0909)	(0.0807)	(0.0876)	(0.1208)	(0.0855)	(0.0807)
N	64,056	18,664	37,945	22,008	25,108	31,306	14,069	24,502	31,653
Sample Cut	None	Lost	Won Last	Won Last 2	Not Covered	Covered Last	Covered Last 2	Pay Week	Non-Pay Week

Notes: Standard errors are clustered at the state-by-season-by-team level. The sample is restricted to game-days where a home team is expected to win (spread less than -4). Columns 2 to 4 further restricts the sample based on whether a team won or lost the previous week. Columns 5 to 7 further restricts the sample based on whether the home team covered the spread (actual score less than the pre-game spread). Columns 8 and 9 further restricts the sample based on Sundays after paydays (1st to 7th and 15th to 21st) or other Sundays. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

In column 1 of Table 3, we present our estimates using the restricted sample and show that our estimated effect is qualitatively similar to our main specifications. Next, we further narrow down our sample based on whether a team lost the last game (column 2), won the last game (column 3), or won the last two games (column 4). We find that the estimated effect becomes larger (0.2025 vs. 0.1257) when a home team is on a winning streak, though we cannot precisely rule out whether these coefficients are statistically different. In columns 5 to 7, we find similar patterns of results when we focus on whether a spread was covered or not.[30] These findings are consistent with the hot hand fallacy or recency bias.

[29] For example, if a home team does not have a bye week or non-Sunday games on a winning streak, that team will have no counterfactuals.

[30] Covering a spread is slightly different from winning because a team who lost but had a point difference smaller than the spread will cover its spread. For example, if a home team's spread was +5 (lose by 5 points), then a home team can cover the spread if they lose by less than 5 points.

An individual who observes their team recently winning may perceive their team as "good", leading them to place more bets on their team and experience stronger adverse reactions to an unexpected loss.

In columns 8 and 9 of Table 3, we restrict our sample based on whether the Sunday coincided with the payday week.[31] Notably, we find that the upset loss coefficients are qualitatively similar between the two groups. However, the upset loss times betting coefficient is large and statistically significant during Sundays following payday (column 8) but small and insignificant on other Sundays (column 9). These findings align with the hypothesis that individuals may allocate more of their funds to sports betting activities when they receive their paychecks, resulting in more pronounced adverse reactions to monetary losses (Dobkin and Puller, 2007; Stephens Jr, 2006).

In Table 4, we examine for heterogeneous treatment effects between in-person vs. mobile betting as well as the size of the betting market. In column 1, we find no evidence of a worsening reaction to upset loss when states legalize in-person betting. Instead, we find that states that legalized mobile betting drive our results. These findings are consistent with the theory that mobile betting has a larger impact on sports gambling participation than in-person betting due to its lower opportunity cost (Hollenbeck et al., 2024).[32] In column 2, we document that states with large handles primarily drive our overall treatment effect. This finding demonstrates that the financial loss resulting from an upset may be one crucial

[31]Since most paydays in the U.S. typically occur on the first day, 15th, and/or last day of the month, we define the "week of payday" as any Sunday falling within the 1st to 7th or 15th to 21st of the month.

[32]Another possibility worth noting is that in-person betting may lead to less time at home during NFL games, which could also explain why the effect on family violence at home is smaller.

mechanism through which individuals react more strongly to unexpected NFL game outcomes.

TABLE 4. Sensitivity of Our Estimates to Type & Size of Betting Markets

	(1)	(2)
Upset Loss	0.0374*	0.0303
	(0.0199)	(0.0198)
Upset Loss* Mobile Betting	0.1667***	
	(0.0611)	
Upset Loss* In-person Betting	-0.0712	
	(0.0679)	
Upset Loss*Betting*$<$ Median per capita handle		0.0748
		(0.0717)
Upset Loss*Betting*$\geq$ Median per capita handle		0.1356**
		(0.0665)
N	301,854	301,854

Notes: Standard errors are clustered at the state-by-season-by-team level. Column (1) examines if the state allows online or in-person sports betting. Column (2) examines heterogeneity by whether a state had larger sports gambling market, as measured by average amount wagered. The median handle used for cutoff is 26 per person. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

2.6 Conclusion

Policymakers and legislatures are advocating for the legalization of sports betting due to its potential tax benefits. Sports betting can generate large tax revenues from displacing the black market, and increase in the number of people who gamble in the state due to the creation of a new market or people deciding to place a bet in that state rather than Nevada. In 2022, total taxes collected among states with legalized sports betting totaled almost \$1.5 billion.[33] On the

[33]This data does not include tax revenue in Montana or Oregon.

other hand, sports gambling legalization can generate negative externalities, such as displaced tax revenues in other gambling markets (Humphreys, 2021; Can et al., 2023).

Another potential negative externality arising from sports gambling legalization is the amplification of emotional cues. In this paper, we are the first to investigate the causal relationship between legalized sports gambling and IPV. Using data from the 2011 to 2022 NIBRS, we document that legalized sports betting amplifies emotional cues, as evidenced by increased IPV when a fan's home team unexpectedly loses.

Although we shed light on some potential consequences of allowing sports gambling, we do not necessarily conclude that states should reverse such legalization. Following a similar line of thought as Humphreys (2021), who suggests that states should carefully examine how legalizing sports betting impacts tax revenues to "mitigate the fiscal consequences of legalization," we suggest that states should examine how they can use state tax revenues from sports betting to mitigate these negative externalities.[34] For instance, states can fund advertising campaigns—perhaps during commercial breaks during sporting events—to raise public awareness of the potential consequences of participating in sports gambling or committing family violence. Moreover, states can also invest their tax revenues

[34]While doing back-of-the-envelope calculations to estimate whether the net social costs from emotional cues outweigh the tax revenue from sports gambling is intriguing, we cannot perform such calculations for several reasons. First, we need to determine the temporal and spatial distribution of IPV instances in the U.S. While the NIBRS provides helpful information about IPV, they are not necessarily nationally representative because some states and counties do not report to the NIBRS. Second, our estimates are specific only to NFL; thus, any derivations of costs will only be specific to September to December. Finally, we do not know the severity of each IPV to measure the total costs. While we find compelling evidence that the number of IPV instances is increasing, we need to figure out whether these IPV instances lead to physical injury or hospitalization.

in domestic violence shelters or hotlines to assist the victims or in counseling services to offer support for potential offenders.

CHAPTER III

IMPROVING PROGRAMMING IN JUVENILE DETENTION: THE IMPACT OF PROJECT SAFE NEIGHBORHOODS YOUTH OUTREACH FORUMS

This chapter is joint work with Assistant Professor Jonathan MV Davis and Walton Hale Hamilton Professor Tracey Meares published in the *Journal of Quantitative Criminology*. Jon brought me on to this paper to contribute to the edits of this paper. I rewrote and edited every section of this paper and contributed to the coding work.

3.1 Introduction

Nearly 50,000 youth, who hail disproportionately from the most economically disadvantaged and socially isolated neighborhoods in America, are incarcerated in a residential facility on a given day (Sickmund et al., 2019). This juvenile incarceration has the potential to serve several policy purposes. Incarceration prevents youth from committing additional crimes while they are detained; the threat of being imprisoned may deter some youth from committing crimes in the first place; and incarceration could serve as an intensive intervention designed to redirect youth to more pro-social activities than crime, like school or formal employment. These potential benefits come at a cost. The state of Illinois spends over $300 per day for every incarcerated youth (Petteruti et al., 2014). But this direct cost may be dwarfed by the long-term costs of incarceration: Youth are unable to attend their regular school while incarcerated. After leaving detention, many individuals have reduced economic opportunities and lower earnings which also affects the individual's family and community (Western and Pettit, 2010). The costs are not limited to economic outcomes: nationally, nearly one in ten incarcerated youth is sexually assaulted while in custody (Beck et al., 2013).

The best evidence to date suggests juvenile incarceration in its current form primarily serves to incapacitate youth. Lee and McCrary (2017) show there is little difference in criminal behavior of youth before and after their 18th birthday, despite a substantial increase in potential punishments in the adult justice system. This suggests the threat of incarceration is not a serious deterrent of juvenile crime. Aizer and Doyle (2015) use differences in the sentencing patterns of randomly assigned judges to identify the impact of juvenile incarceration on adult outcomes. They find juvenile incarceration reduces the probability of graduation and increases the probability of incarceration later in life. This suggests juvenile incarceration may be pushing youth further into crime rather than redirecting them to more pro-social activities. Since incapacitation is the shortest lived of the potential benefits of juvenile detention, these results suggest juvenile incarceration in its current form may not be an efficient policy. This raises the question: are there aspects of juvenile detention that can be improved?

This study addresses the narrower question of whether a Youth Outreach Forums program that has been run in the Cook County Juvenile Temporary Detention Center (JTDC) by the Northern Illinois Project Safe Neighborhoods (PSN) Task Force since 2015 can improve the outcomes of detained youth. The core of the program is a series of four forums that were usually held over four consecutive days during school breaks. Each forum provides participants the opportunity to have a structured discussion in as neutral a setting as possible with authority figures in their lives, including family members, teachers, and law enforcement agents, as well as individuals with insights about the consequences of crime, including community members and ex-offenders. The program was inspired by quasi-experimental evidence that a similar program serving adult probationers led to a significant reduction in homicide rates (Papachristos et al., 2007).

The theoretical basis for the program is the idea that most people voluntarily comply with laws and rules they consider legitimate (Tyler, 2006a,b; Tyler and Trinkner, 2018). According to procedural justice theory, this legitimacy can be established by treating individuals with dignity and respect and giving them a voice in the interaction and by demonstrating that the authority figure is neutral and trustworthy. Bottoms and Tankebe (2012) go further and explain that legitimacy is dialectic in nature, requiring both the audience and the power-holders to agree on the legitimacy. The hope is giving youth the opportunity to have an open discussion with the authority figures in their lives about the logic behind the rules that are in place and the consequences of breaking those rules will better establish this "audience legitimacy" by enhancing perceptions of the procedural justice of their interactions with the legal authorities with whom they are dealing (Tyler and Huo, 2002). Although specifically motivated by procedural justice theory, the program can more generally be considered to be a focused deterrence strategy. Braga and Kennedy (2021, p. 3) explain that focused deterrence strategies "seek to change offender behavior by understanding underlying violence-producing dynamics and conditions that sustain recurring violence problems, and implementing a blended set of law enforcement, informal social control, and social service actions." In this spirit, the program leverages many elements that could all reduce crime. For example, the program could reduce crime by making the long-term consequences of youth's actions more salient, developing participants' skills, introducing the youth to positive adult mentors, or connecting the youth with community support.

To credibly measure the impact of the program on youths' outcomes, the program was implemented in the JTDC as a randomized controlled trial in 2015 and 2016. The research team randomly selected groups of youth to be invited to voluntarily participate in the program using a cluster-randomized design. Because the treatment is randomly assigned, any differences in outcomes between the treatment and control groups can be credibly attributed to the program.

Our preferred estimates indicate that the program caused a 20 percent reduction in the number of new spells at the JTDC in the eight months after random assignment and reduced total arrests by 18 percent in the first year after random assignment. While both of these impacts are somewhat imprecisely estimated, the reduction in total arrests is driven by statistically significant 43 and 40 percent reductions in arrests for violent and drug crime, respectively, and a large but less precisely estimated 30 percent reduction in arrests for property crime. These correspond to very valuable and proportionally large reductions in the social costs of crime. There is a small and insignificant reduction in other crime. Our estimates indicate proportionally large, but imprecisely estimated, increases in school attendance and grade point averages (GPAs). However, these increases are relative to very low baseline levels of school attachment. These results provide suggestive evidence that offering similar programming to incarcerated youth could reduce future criminal behavior and increase attachment to school.

These results contribute to a growing literature that demonstrates the potential of focused deterrence strategies to reduce crime. Braga et al. (2018) conduct a systematic review of 24 quasi-experimental evaluations of focused deterrence strategies. They note that 22 of the 24 studies in their review found beneficial impacts on crime with the other 2 studies finding basically no difference between the treatment and comparison conditions. Notably, none of the studies in their review used a randomized experimental design. Our study helps fill this gap in the literature.

Our results are also an encouraging contribution to the small experimental literature evaluating whether particular programs can improve outcomes for incarcerated individuals. Farrington and Welsh (2005) review 14 experiments evaluating correctional programs. Two of these studies found "scared straight" programs increase recidivism and four found juvenile bootcamps have no effect or have a detrimental effect on re-offending. The evidence from eight experiments evaluating therapeutic programs offered to inmates are more encouraging, but the effects are usually statistically insignificant. Grommon et al. (2018) experimentally evaluate a dog-training program offered to youth in a midwestern juvenile detention center and find that it has no effect on self-reported psychosocial measures, including self-esteem, empathy, optimism, pessimism, compassion, and social competence. Heller et al. (2017) leverage a natural experiment in the JTDC, the same setting as the present study, and find residents who were randomly selected to receive a set of reforms that increased the educational requirements for staff, instituted a new system of rewards for youth, and introduced youth to cognitive behavioral therapy techniques for managing conflict and bad behavior were 21 percent less likely to return to the JTDC. Programs that seek to keep youth out of the juvenile justice system through restorative justice conferencing also have shown mixed results. Recently, Kimbrell et al. (2023) find promising results through a meta-analysis of both experimental and quasi-experimental research designs that restorative justice programs are more impactful than other conventional methods. Shem-Tov et al. (2021) find that participation in the Make-it-Right (MIR) restorative justice program in San Francisco, which worked with youth ages 13 to 17 who committed a felony, reduced recidivism. In contrast, Vooren et al. (2022) find that participation in Halt, a restorative justice program for youth in the Netherlands, increased the probability of recidivism one year after the program and decreased educational attainment.

Finally, our study provides lessons on running experiments in detention centers. In the first year of the study, random assignment occurred about two weeks before the program was scheduled to begin in order to give JTDC staff time to prepare for the program. This had the unintended consequence of making treatment uncorrelated with program participation because of youth's typically short spells in the JTDC. Many youth who were assigned to the treatment group were no longer at the JTDC when the program was run. After observing the first year results, the research team and program administrators made adjustments to the randomization protocol in order to increase the relevance of random assignment in the second year of the experiment, including moving it much closer to the start of the program. Given that these adjustments were effective, our preferred estimates are from the second year of the study. Because the first year of the study is not informative about treatment effects, our estimates are less precise than expected in the experimental design. Given the imprecision, we emphasize effect sizes and confidence intervals in addition to p-values. Recent debates in the statistics literature on the value of hypothesis testing have emphasized that a focus on effect sizes and confidence intervals may better address whether a hypothesis is correct than the traditional focus on p-values (Wasserstein and Lazar, 2016; Benjamin et al., 2018).

The remainder of this paper is organized as follows. Section 3.2 provides details about the program and its development. Section 3.3 explains the experimental design and the details of the analysis. Section 3.4 provides descriptive statistics on the youth in our sample. Section 3.5 presents the program's impact on youth's post-randomization juvenile detention, arrests, and schooling. Section 3.6 concludes.

3.2 Context and Intervention

Antecedents of the Intervention

The Project Safe Neighborhoods (PSN) initiative was launched in 2001 by the United States Department of Justice to reduce gun violence (McGarrell et al., 2009). The initiative provides each of the 94 U.S. Attorney districts with financial resources, training, and a framework to support local solutions to gun violence. Each region formed a PSN taskforce that included representatives of the U.S. Attorney's Office, other federal, state, and local law enforcement agencies, members of the broader community, including representatives from local governments, schools, and social service agencies, and researchers. While the PSN taskforces could develop solutions developed to their own local context, each taskforce was provided with training on intensively prosecuting violent gun crime, interrupting the supply of illegal guns, holding offender deterrence meetings, and offering support services to help shift individuals out of crime. In the first seven years of the program, Congress allocated nearly 3 billion dollars to support the PSN initiative (McGarrell et al., 2009). In 2022, the Office of Justice programs awarded about 17.5 million in PSN grants.[1]

The PSN initiative was inspired by the perceived success of Richmond's Project Exile and Boston's Operation Ceasefire. Richmond's Project Exile was a joint effort between law enforcement in Richmond, Virginia and the regional U.S. Attorney's office to prosecute all drug and domestic violence cases involving guns and all felon-in-possession-of-a-firearm cases in federal courts (Raphael and Ludwig, 2003). Federal sentencing guidelines for gun crime were more severe than those in Virginia when the policy was in operation and would normally result in sentences in out-of-state prison. An advertising campaign was used to inform the community and potential offenders about the increased certainty of enforcement. The policy took effect in February 1997. A 35 percent decline in the homicide rate and a 37 percent decline in the gun homicide rate between 1997 and 1998 contributed to the perception of Project Exile's success. However, Raphael and Ludwig (2003) show that most of this decline was driven by an unusually high homicide rate in 1997. They show that more credible research designs suggest the program had little to no effect.

Boston's Operation Ceasefire was a partnership between the Boston Police Department, the Bureau of Alcohol, Tobacco, and Firearms, the U.S. Attorney, the Suffolk County District Attorney, the Massachusetts Department of Probation, the City of Boston, the Massachusetts Department of Parole, social service workers, and researchers (Kennedy, 1997). The intervention was motivated by the insights that the justice system was not effectively deterring crime because punishments were uncertain and insubstantial and that a small number of individuals are

[1] https://data.ojp.usdoj.gov/stories/s/Office-of-Justice-Programs-Funding/

responsible for a large share of crime. The program aimed to more effectively deter violent crime using a "focused deterrence strategy." Legal authorities would respond to violent crime by "pulling every lever" to impose additional costs on committing crime. A violent crime will bring heightened law enforcement attention to all crimes, including to more minor violations, like probation violations, loitering, or drinking in public. This new enforcement strategy was communicated at formal meetings with gangs and gang members and presentations at schools and juvenile detention centers, and by word of mouth through social workers and law enforcement agents.

Braga et al. (2001) measure the impact of Operation Ceasefire by comparing the number of youth homicides in Boston in the years before and after the intervention took place and also by comparing trends in Boston's crime outcomes to analogous trends in other cities. The pre-post comparison shows that Operation Ceasefire led to a large proportional reduction in youth homicides. The comparison with other cities showed that Boston's crime reduction was relatively large compared to typical changes during that time period, but several other cities saw similar declines.

This type of focused deterrence strategy has also been used to address drug markets. The High Point Intervention in High Point, North Carolina addressed drug markets by targeting a particular drug market, arresting any violent drug dealers in that market, and having non-violent drug dealers attend a "call-in" meeting where community members, social service providers, former offenders, and law enforcement tell the drug dealers that they are valued in the community and that the drug dealing needs to stop (Kennedy, 2009). The non-violent dealers are informed that law enforcement has developed a case against them, but the case will not be brought unless they re-offend. Using quasi-experimental methods, Corsaro et al. (2012) find that the High Point Intervention led to modest violent crime reductions, especially in the first two areas that were targeted.

In addition to addressing group crime dynamics, like gang violence or drug markets, focused deterrence strategies have also been used to address recidivism of individual offenders. In Chicago, a key element of the PSN strategy was a series of "Offender Notification Forums" (Papachristos et al., 2007). Offenders with a history of gun violence were asked to attend an hour-long forum about the consequences of gun violence and how to transition away from crime. The meetings began with a 15 minute discussion with law enforcement about local and federal laws regulating gun violence and the consequences of violating these laws. Next, an ex-offender would lead a discussion of how he had successfully avoided reoffending. The forum would conclude with a discussion with social service providers in the community about what resources are available to support their success. These forums were designed with the goal of establishing legitimacy. Meetings were held at neutral locations, like parks or libraries. All participants would sit around a table on a level playing field. The PSN taskforce complemented these forums with federal prosecutions of gun crime, gun recovery efforts, and community outreach. Papachristos et al. (2007) find that these efforts caused a large proportional reduction in homicides in targeted police districts using quasi-experimental methods.

The Northern District of Illinois's PSN task force also used group violence reduction strategies. Papachristos and Kirk (2015) discuss Chicago's 'call-in" program. These meetings brought together groups of gang members who had been identified by the police as being involved in potentially violent conflicts, police leadership, and community members. The police informed the gang members that the next shooting they were involved in would get the full attention of law enforcement. Community members emphasized that the gang members were part of their community but that the violence had to stop.

Focused deterrence strategies are now an important component of crime prevention policies. As demonstrated by the above examples, focused deterrence strategies generally follow the same framework. An interagency task force is brought together to address a particular crime problem. The crime is then addressed with special enforcement efforts that are complemented with an outreach campaign informing potential offenders about the enforcement, the impact of crime on the community, and the services available to support their success (Kennedy, 2006). Braga and Kennedy (2021) provide a more detailed overview of the theory, development, and applications of focused deterrence strategies. They highlight that the effectiveness of focused deterrence is supported by its "strong logic model" that creates several mechanisms that could all help reduce crime. First, focused deterrence more effectively generates a deterrent effect by combining enhanced punishments with clear communication to potential offenders about the change in enforcement. Second, focused deterrence promotes legitimacy and procedural justice by encouraging respectful, direct communication between legal authorities and potential offenders (Papachristos et al., 2007; Braga and Kennedy, 2021; Bottoms and Tankebe, 2012). Third, the participation of community members and social service providers reinforces informal social control. They actively highlight the negative impact of crime on the community, increasing awareness and accountability. Furthermore, by connecting potential offenders with services that not only support their personal success, but also underscore the community's investment in their well-being.

Braga et al. (2018) conduct a meta-analysis of the evaluations of focused deterrence strategies. They identified 24 quasi-experimental evaluations. The evidence presented in these studies suggests that focused deterrence strategies causes a moderately large reduction in crime outcomes. However, they find that the effects are smaller, but still beneficial, in studies using research designs that they identify to be more credible. Their review did not identify any evidence from randomized controlled trials.

The Program

The PSN Youth Outreach Forums program was developed by the Northern District of Illinois' PSN Task Force and was inspired by adult outreach forums studied by Papachristos et al. (2007). Like the adult forums, the youth forums can be thought of as an individual focused deterrence strategy because the forum participants are not necessarily members of a shared group. With that said, the youth intervention studied here and the adult intervention studied by Papachristos et al. (2007) have some key differences. The youth program is implemented while participants are incarcerated in a concentrated four-day setting rather than a long-term setting when they are released. The primary component of the program is a series of four forums which provide participants the opportunity to have a structured discussion in as neutral a setting as possible with authority figures in their lives, including family members, teachers, and law enforcement agents, as well as individuals with insights about the consequences of crime, including community members and ex-offenders who have moved away from crime. Youth sit in a circle with the speakers in order to put the youth and speakers on as level a footing as possible consistent with the tenets of the theory, which emphasize, among other factors, the importance of recognizing individual dignity.

The program was managed by the Northern District of Illinois' U.S. Attorney's Office and implemented by the PSN Task Force in partnership with the Cook County Juvenile Detention Center. The Northern District of Illinois' PSN Task Force implements many programs reaching both youth and adults. All of the forums were facilitated by a member of the PSN Task Force from the Chicago Police Department.[2]

The curriculum consists of four forums which were typically held over four consecutive days during school breaks. The forums were offered over breaks because residents attend school based on a typical school year calendar at the JTDC's Nancy B. Jefferson school. Figure 1 shows the daily population of youth in JTDC and the date when youth were randomized into the program. During school breaks, there may not have been as much regular programming, so this program also provided additional programming to youth in JTDC.

The first forum is "Walking the Walk."[3] In this forum, ex-offenders talk to youth about the consequences of their criminal life. This includes a presentation by ex-gang members affiliated with the organization "In My Shoes" who have permanent disabilities due to violence resulting from their gang involvement[4] and an adult who spent time in the JTDC as a teenager who spent most of his

[2]The program facilitator explained to the youth that as Deputy Director for Community Engagement she was an employee of the police department but not a police officer.

[3]The order of the forums changed over the course of the experiment, so we describe them in the order that they were implemented in the second year of the experiment here.

[4]This discussion is graphic. While observing one session, one coauthor nearly had to leave the room after becoming flushed and lightheaded.

adult life in prison because of his continued involvement with the criminal justice system as a young adult. He emphasizes that he felt like he was "getting away with it" as a teenager until he was the subject of a federal indictment that resulted in more than two decades in prison. These presentations aim to reinforce how participants' lives may be adversely affected by ongoing involvement in crime and continued contact with the criminal justice system. The "In My Shoes" presentation emphasizes the potential non-legal consequences of being involved in crime (Braga and Kennedy, 2021). The second presentation enhances the salience of potentially escalating criminal consequences. Both presentations contribute to deterrence by emphasizing the costs associated with criminal behavior.

The second forum is titled "Mutual Respect with the Law." This forum is designed to increase youth's trust in law enforcement and establish the legitimacy of law enforcement while decreasing youth's cynicism about the law. In the forum, youth have a neutral conversation with uniformed police officers with the goal of showing the participants that police officers are regular people, explaining why police behave the way they do, and reinforcing the consequences of continued involvement with the justice system. All of the officers were members of the Chicago Police Department's command staff. Most of the youth recognized that the gold stars on the officers' uniforms signaled that they were high-ranking officers. This is intentional because it gives the youth the opportunity to talk about their negative interactions with police officers with decision makers who can also speak with authority about the department's policies and procedures. As Bottoms and Tankebe (2012) discuss, misconduct by authority figures undermines legitimacy and creates resentment. The discussion of how police are themselves held accountable directly addresses this source of illegitimacy. The presence of command staff at the forum could also have a deterrent effect by emphasizing the importance of the forums and showing the youth that they are under scrutiny (Braga and Kennedy, 2021). Anecdotally, this forum had a big impact on some participants. Many youth would begin the forum defensively and would not speak to the officers, but by the end of the session, most of the youth were willing to talk to the officers and even shake hands.

The third forum is titled “Community and a Shared Moral Voice.” The goal of this session is to help youth understand community norms, values, and morals. This is accomplished in two ways. First, community members are invited to have a conversation with youth about how crime affects the community. This promotes legitimacy by establishing that rules are supported by a set of shared beliefs (Bottoms and Tankebe, 2012). Second, youth are introduced to resources available in the community that are designed to help them continue to make better decisions once they are released. These social supports could reduce crime for a number of reasons. For example, they could directly solve a problem (i.e. helping pay bills) that the youth may have tried to solve by committing a crime (i.e. selling drugs). As another example, they could also enhance youth’s sense of belonging in their community.

The final forum’s focus changed over the course of the experiment. It was originally called “Hope Outside the Law.” The session aimed to help youth develop more positive relationships with non-legal authorities, their parents and teachers in particular, understand the importance of behavioral standards in non-legal environments, and increase youth’s future orientation. This session was designed to be split evenly between parent-child relationships and teacher-student relationships. Youth would be introduced to a school re-engagement specialist employed by the program at the end of this session. Reengagement specialists work with the youth to find the school within Chicago Public Schools that will best meet their needs and maximize their chances of graduating. In practice, it was difficult to implement "Hope Outside the Law" as intended because it required parents to make a sometimes long trip to the JTDC to participate and teachers to volunteer their time on their off days. In one case, no parents showed up and that portion of the session needed to be improvised. Even when some parents did attend, the program facilitators worried that this emphasized the absence of caring adults in some of the other children's own lives.

Given the challenges in implementing the "Hope Outside the Law" forum, it was replaced by a new forum focused on conflict resolution strategies in the second year of the study. This session is built around cognitive behavioral therapy principles and gives youth the opportunity to practice conflict resolution strategies in the context of interactive games. For example, in one activity, a single youth is given instructions to do a non-verbal action. The other kids are then asked to discuss what message the youth is sending. The facilitator would highlight that the wide variety of responses highlights that people interpret things in different ways. Recognizing that they may mistakenly interpret something as aggressive or antagonistic could help youth avoid entering into a conflict unnecessarily. Similar cognitive behavioral therapy programming has been convincingly shown to reduce youth's future criminal behavior (Heller et al., 2017).

Although the vast majority of youth in the study are boys, several adjustments are made when girls participate. In particular, the "Walking the Walk" forum is adjusted to include female ex-offenders and the program "In Her Shoes" which focuses on domestic violence issues. The "Mutual Respect with the Law" forum includes female command staff. Inviting female command staff not only gives the girls the opportunity to have the same discussion about their rights and responsibilities while dealing with law enforcement, but also gives them exposure to women in positions of power.

The forums were funded by a $500,000 Bureau of Justice Assistance grant to support PSN initiatives in the Northern Illinois District. Much of this funding supported the cost of personnel required to develop and coordinate the youth forums program. The marginal cost of running each forum is much less, only a couple hundred dollars, given that many of the presenters volunteered their time.

The Facility

The Juvenile Temporary Detention Center (JTDC) in Cook County, Illinois is the largest juvenile detention center in the United States. More than 80 percent of youth admitted to the JTDC are between 15 and 17 years old, but youth as young as 12 and as old as 20 were admitted during our sample window (March 1st, 2015 through April 30th, 2017). The average youth in our sample is almost exactly 16 years old on the day of their admission.

There were 2,622 admissions to the JTDC by 1,718 unique youth in the year starting July 1st, 2015. On average, there were 298 youth housed at the JTDC on any given day during this period. Fig. 6 shows the daily population based on the juvenile court records shared with the research team covering all admissions between March 1st, 2015 and April 30th, 2017. There is a clear seasonality to the population, with the population highest in the summer and lowest in the winter. Most stays at the JTDC are relatively short. The median completed spell in our data is 15 days and the interquartile range is 3 to 31 days. The subset of youth who are being housed in the JTDC until they are old enough to transfer to the adult court system are classified as "automatic transfers" (AT). These youth may have substantially longer stays as they are awaiting transfer to the adult court system.

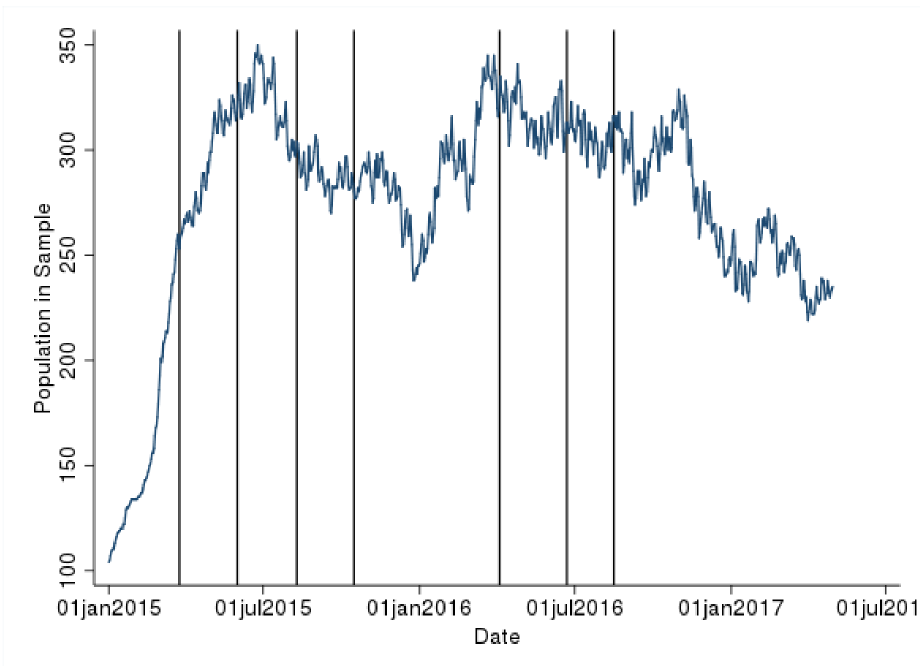


FIGURE 6. Daily Population in the JTDC

Notes. This figure shows the daily population of youth admitted between January 1, 2015 and April 30, 2017. The vertical bars indicate the dates of each lottery.

Residents of the JTDC are organized into a series of residential "pods" which typically house 8 to 12 youth. With few exceptions, youth participate in nearly all offered activities together with the other youth in their pod and rarely interact with youth from other pods. Although youth may be offered activities within their pod, participation is voluntary. Pods are segregated by gender and whether or not a youth is an AT.

3.3 Experimental Design and Analysis

The experiment includes seven rounds of forums that were run between March 2015 and August 2016. For each round of forums, JTDC administrators provided the research team with a list of pods eligible to participate in the forums. The research team randomly selected a pre-specified number of pods using a block randomized design. Typically, three or four pods were selected for treatment in each round of the lottery so that each session of the forum could be held on the same day. Blocks were defined by lottery, AT status, and gender. Randomization was at the pod level because, with few exceptions, youth complete all activities together with their pods.

Table 5 shows the number of clusters included and selected for treatment across all lotteries and separately for the first and second year of the experiment. A cluster is defined as a pod-lottery combination. Most pods were included in multiple lotteries, but because most youth spend less than a month in the JTDC, only a small number of youth were included in multiple lotteries. The experiment includes a total of 86 clusters of which 26 were selected to receive the program. The first and second year of the study include 56 and 30 total clusters of which 17 and 9 were selected for the treatment group, respectively.

TABLE 5. Treatment Assignment Details

	Number of Clusters	Number Assigned to Treatment
All Lotteries	86	26
Year One	56	17
Year Two	30	9

Notes: Treatment was randomly assigned at the residential pod level. This table shows the number of eligible pods and the number of pods selected for treatment in the entire study and in the first and second year of the study.

Individual youth are considered treated if, on the day of the lottery, they resided in a pod which was selected for treatment. Youth are part of the control group if, on the day of the lottery, they reside in a pod that was not selected for treatment. To be conservative, we do not include individuals who move into a treatment or control pod after the day of randomization in the study because knowledge of a pod's treatment status may have affected how youth were assigned to pods.

Table 6 shows the number of youth participating in at least one forum by treatment status and year of the study. In order to protect youth from self-incrimination, youth were asked not to share their names or any details about their cases during the forums. However, youth were asked to submit a notecard with their name to the JTDC's staff in order to receive a certificate for completing the program. We measure participation by linking the names on these notecards with our analysis sample.[5] Youth were not required to submit a card or to write their real name.[6] A total of 16 participants who submitted a notecard could not be matched. We do not have data on how many youth did not submit a notecard. Therefore, our participation measure is a lower bound on actual participation. Participation information is available for six of the seven lotteries. Participation data is not available for the first forum series.

In the first year of the study, the lottery was conducted by the research team about two weeks before the scheduled start of the forums in order to give the JTDC's staff time to prepare for the program. Unfortunately, few youth living

[5] We match the participation data to our court records by first running a fuzzy match, allowing for minor differences in first and last names. We then manually look for matches for any unmatched participants.

[6] Some youth opted to write the name of fictional movie characters.

TABLE 6. Treatment Status and Participation by Year

A. Participants by Treatment Status and Year

	Treatment			Control		
	Non-Participant	Participant	Participation Rate	Non-Participant	Participant	Participation Rate
All Years	139	41	23%	353	16	4%
Year One	91	11	11%	184	10	5%
Year Two	48	30	38%	169	6	3%

B. Impact of Treatment Assignment on Program Participation

	Pooled (Lotteries 2 through 7)	Year One	Year Two
ITT	0.164***	-0.02	0.347***
	(0.07)	(0.08)	(0.09)
Control Mean	0.04	0.05	0.03
Number of Observations	549	296	253
Number of Pods	74	44	30

Notes: Panel A shows the number of youth in the treatment and control groups who did and did not participate in the program and the implied participation rate. Panel B shows the impact of being randomly assigned to treatment on the probability of actually participating in the program controlling for baseline covariates and block fixed effects. In order to protect youth from self-incrimination, youth were asked not to share their names or any details about their cases during the forums. However, youth were asked to submit a notecard with their name to the JTDC's staff in order to receive a certificate for completing the program. We measure participation by linking the names on these notecards with our analysis sample. Participation information is available for six of the seven lotteries. Participation data is not available for the first forum series, so youth in this randomization are not included in this table. Our measure of participation is a lower bound for participation because youth could choose not to turn in a notecard or they could turn in a notecard with a fake name. Stars indicate: * $p<0.1$, ** $p<0.05$, *** $p<0.01$.

in the pod on the day of the lottery were still in that pod when the program began, either because they were released or moved pods. As a result, treatment group youth were only 6 percentage points (pp) more likely to participate in the program than control group youth in the first year. After controlling for individual covariates and randomization block fixed effects (see subsection for details), the first-stage impact of treatment status on participation shown in Panel B is -0.02 with a standard error (SE) of 0.08 and p-value on the null hypothesis that the impact is zero of 0.81. Therefore, treatment assignment is not a relevant predictor of program participation in the first year of the study.

In the second year, after observing the low participation rates and substantial control cross-over in the first year, we ran the lottery just a few days before the program was scheduled to begin.[7] This change solved the problem: In the second year of the study, treatment group youth were 35 pp more likely to have participated in the program than control group youth (38 percent versus 3 percent). The first-stage coefficient in the second year is 0.35 (SE=0.09, p<0.01).

Because random assignment is not a relevant predictor of program participation in the first year, our preferred analysis focuses on the second year of the program.[8]

Data

Our sample is derived from Juvenile Court Data on the housing history of residents of the JTDC. This data covers all youth who were admitted to the JTDC between March 1st, 2015 and April 30th, 2017. Each record includes the youth's

[7] This change occurred for lotteries after January 1st, 2016.

[8] Appendix Tables A.4 and A.5 show results for the full sample and for the first year, respectively, for completeness and transparency.

name, date of birth, and admission and release dates to the JTDC and to each pod. A youth is included in our analysis sample if he or she was assigned to a pod that was included in a lottery on the date the lottery was run. We also use these juvenile court data to measure details about youth's pre- and post-randomization spells in the JTDC.

We measure arrests using Illinois State Police (ISP) arrest records provided by the Illinois Criminal Justice Information Authority (ICJIA) which combine police records from departments across the state. Youth are matched to these records by ICJIA using their names and dates of birth. The data cover both juvenile and adult arrests from 2001 through September 2017, a little over a year after the final lottery was run. Each record includes information about the arrested individual as well as the date of the incident and a description of the charge. We use the charge description to classify whether each arrest is for a violent, property, drug, or other crime. 98 percent of youth are matched to a pre-randomization arrest. In principle all youth should have a prior arrest, but some arrests may have been missed because of substantial typos in names or birth dates.

We measure schooling histories and outcomes using student-level records from Chicago Public Schools. These data include information on youth's enrollment status, attendance, and grades. We match residents to these records using fuzzy matching techniques on names and dates of birth.[9] 98 percent of residents in our sample are matched to a Chicago Public Schools record.

To improve our statistical power, we also report treatment effects on a "standardized index" of detention, arrest, and schooling outcomes. To generate

[9]Specifically, we measure the quality of first and last name matches using the Jaro-Winkler distance between the strings and allow for single digit typos in dates of birth.

this index, we first standardize each outcome using the control group mean and standard deviation. Then we average the standardized outcomes.

Estimation and Inference

Our sample includes one observation for each unique youth-lottery combination for a total of 626 observations on 609 youth. Of these 609 youth, 594 were included in one lottery, 13 were included in two lotteries, and 2 were included in 3 lotteries.

Our analysis is based on the following individual-level panel regression:

$$Y_{ipst} = \gamma T_{pt} + X_{i,t-1}\beta + \alpha_{st} + \varepsilon_{ipst}, \tag{3.1}$$

where i indexes youth, p indexes pods, s indicates whether the pod housed juvenile male, AT, or female residents, and t tracks the date of the lottery. Y_{ipst} is an outcome for resident i living in pod p of type s that was included in the lottery at time t. T_{pt} is an indicator for whether or not pod p was assigned to the treatment group on date t. This indicator equals zero if a youth lived in a pod which participated in the lottery but was not selected for the treatment group.

Randomization blocks are defined by whether the pod housed juvenile males, AT males, or female residents and the date of the lottery. α_{st} is a randomization block fixed effect. $X_{i,t-1}$ is a vector of pre-randomization covariates, including controls for age (in years) on the day of the lottery, spells in the JTDC since March 1st, 2015, length of the current spell in the JTDC on the day of the lottery, number of arrests for violent, property, drug, and other crimes in the year before the lottery, indicators for having zero violent, property, drug, or other arrests, days attended in the school year before the lottery at non-prison schools, GPA in the semester before the lottery, the number of prior lotteries the youth was included in, and the number of prior forum series the youth attended. Separate indicator variables for missing any of these control variables are also included.[10] These baseline controls are not necessary for identification because T_{pt} is random conditional on randomization block. However, their inclusion improves our statistical power.

The coefficient γ is interpretable as the intent-to-treat (ITT) effect of the program as it is the impact of being randomly assigned to the treatment group, regardless of whether or not the resident actually participates in the program. This is a policy-relevant parameter as it is informative about the average impact of the program given the actual participation rates associated with the program.[11]

Due to the implementation challenges in the first year of the study, our preferred estimates are from the second year of the program. These estimates come from the following adapted version of Equation 3.1:

$$Y_{ipst} = \gamma_1 T_{pt} \times (t \leq 4) + \gamma_2 T_{pt} \times (t \geq 5) + X_{i,t-1}\beta + \alpha_{st} + \varepsilon_{ipst}. \tag{3.2}$$

γ_1 and γ_2 are the intent-to-treat effects in the first and second years of the study, respectively. γ_2 is our preferred estimate. We use a pooled regression that includes

[10] Missing values are imputed as zero.

[11] We focus on the ITT instead of the direct impact of participation because participation is potentially quite noisily measured because of non-response, providing fake names, etc. We estimate that random assignment increased participation by 34.7 pp in the second year of the study so the local average treatment effect (LATE) would be 2.88 times larger than the ITT (LATE = ITT/.347) (Angrist et al., 1996). If we underestimate participation, however, this adjustment will be too large.

the first year of the study, instead of estimating Equation 3.1 separately by study year, to more precisely estimate the coefficients on the control variables.

Standard errors are clustered by both individual and pod-lottery cluster. We implement this two-way clustering using the method of Cameron and Miller (2015).

3.4 Descriptive Statistics

Table 7 shows descriptive statistics of the control group and estimates of treatment-control differences separately for the first and second year of the study. We show results separately by year, as we will focus on the second year of the study in the remainder of the analysis because of the compliance issues in the first year.

An average resident in the control group in the first year of the study was 15.8 years old with 1.4 spells in the JTDC at baseline, including their current spell at the time of the lottery, who had been incarcerated in the JTDC for about 25 days. In the second year of the study, an average resident in the control group was 16.2 years old, had 2.8 spells in the JTDC, and had been in the JTDC for 28 days. Not surprisingly, nearly all youth have a pre-randomization arrest in the year before the study (99 percent in year one, 97 percent in year two).[12] On average, control group youth in the first year of the study had 1.74 arrests for violent crime, 1.15 arrests for property crime, 1.36 arrests for drug crime, and 4.93 arrests for other crime in the year before the study. In the second year, control youth had 2.29, 1.42,

[12]Youth may not have any baseline arrests because of discrepancies in their name or birth date between the court records and Illinois State Police arrest records that prevent them from being matched.

TABLE 7. Baseline Balance

	Year One				Year Two			
	N	Control Mean	Treatment Coefficient	SE	N	Control Mean	Treatment Coefficient	SE
A. Demographics								
Age at Program Start	373	15.80	0.05	(0.32)	252	16.16	-0.05	(0.19)
B. Juvenile Detention								
Number of Pre-Lottery Spells in JTDC	373	1.43	0.04	(0.10)	253	2.81	0.21	(0.26)
Days Since Admission	373	25.38	1.06	(1.82)	253	27.84	-6.95**	(3.29)
Number of Prior Lotteries	373	0.03	0.009	(0.01)	253	0.05	-0.015	(0.03)
Number of Prior Forums	373	-			253	0.01	-0.005	(0.01)
C. Arrests in Year Before Lottery								
Has Baseline Arrest	373	0.99	-0.01	(0.00)	253	0.97	0.01	(0.02)
# Arrests: Violent	373	1.74	0.42*	(0.25)	253	2.29	-0.16	(0.32)
# Arrests: Property	373	1.15	-0.10	(0.20)	253	1.42	-0.18	(0.25)
# Arrests: Drug	373	1.36	-0.38*	(0.21)	253	1.29	0.04	(0.41)
# Arrests: Other	373	4.93	-0.59	(0.50)	253	5.19	-0.21	(0.57)
D. Prior School Year Academics								
In CPS Data	373	0.99	-0.01	(0.01)	253	0.97	-0.01	(0.03)
Attended Any Days	373	0.75	-0.07	(0.04)	253	0.69	-0.02	(0.08)
Days Attended (including 0s)	373	61.09	1.35	(7.12)	253	44.4	7.26	(7.61)
GPA in Prior Semester (if available)	243	1.39	-0.08	(0.12)	110	1.52	0.12	(0.17)
Joint Significance Test	F(17,55)=1.47, P(F¿1.47)=0.14				F(18,29)=1.06, P(F¿1.06)=0.43			

Notes: The sample includes one observation for each lottery that a youth was included in. The full sample includes 626 observations, 373 are from the first year of the study and the remaining 253 are from the second year of the study. Balance test shows treatment coefficient and standard error from a regression of each characteristic on a treatment indicator and block fixed effects. Random assignment was stratified by whether a pod housed juvenile males, automatic transfer males (who will automatically be transferred to the adult justice system), or females. 89.6 percent of observations are from juvenile male pods, 5.8 percent are from male automatic transfer pods, and the remaining 4.6 percent are from mixed female pods. Stars indicate: * $p<0.1$, ** $p<0.05$, *** $p<0.01$.

1.29, and 5.19 arrests for violent, property, drug, and other crimes.[13]

Nearly all control group youth (99 percent in year one, 97 percent in year two) are matched to a Chicago Public Schools record, but these youth are quite disconnected from school on average. About 75 percent of youth in the first year of the study and 69 percent of youth in the second year of the study attended at least one day of school at a non-prison school in the school year before the program, but they only attended 61 and 44 days on average in the first and second year of the study. Among the group that attended at least one day, youth attended about 81 and 64 days on average in the first and second year of the study, respectively. Given that the school year is 178 days long, this indicates that even youth who attend some school still miss more than half of the school year. Among the subsample of youth with nonmissing grades, the average GPA in the semester before the lottery was a 1.4 in the first year of the study and a 1.5 in the second year of the study, which are between a C and a D average.

Among the 27 treatment-control comparisons in Table 7 one is statistically significant at the 5 percent level and two are significant at the 10 percent level. This is almost exactly what we would expect from chance variation. We test whether the covariates are jointly insignificant by regressing treatment on our full set of baseline covariates, controlling for block fixed effects, and running an F-test of the null hypothesis that the covariates are jointly insignificant. The p-value of the joint test is 0.14 in the first year and 0.43 in the second year. Therefore, as expected because treatment was randomly assigned, there is little evidence of

[13]Gender is not shown in the table since it is collinear with randomization block. Gender is observed for 622 of the 626 observations. 97 percent of the sample is reported as male in either the JTDC's records or in the Chicago Public Schools records. All females appear in the first year of the study.

observable differences between the treatment and control group in either year of the study.

3.5 Results

This section describes the program's impact on juvenile incarceration in the JTDC, arrests, and schooling. Table 8 shows impacts for the second year of the experiment. The results in Table 8 are our preferred estimates because, as discussed in Section 3.3, random assignment was not a relevant predictor of program participation in the first year of the experiment. As a result, we do not expect the program to have any impact, positive or negative, in the first year of the experiment. We include the pooled results and first-year results in Appendix Tables A.4 and A.5, respectively, for completeness and transparency.

Impact on Detention Outcomes

We begin by looking at the impact of being randomly invited to participate in the program on youth's post-randomization juvenile detention outcomes.

Although we are most interested in youth's behavior after being released from the JTDC, we first explore whether the program affected the time until youth are released from the JTDC. Because youth cannot be arrested and cannot enroll at non-prison schools while incarcerated, any impact on time until release could affect our interpretation of the program's impact on other outcomes. Figure 7 shows the Kaplan-Meier survival function for the time until release by treatment status, or the proportion of youth in each group who had not been released after D days. The treatment and control group curves look very similar, which suggests that the program did not impact time until release.

TABLE 8. Intent-to-Treat Effect Estimates

			95% Confidence Interval		Control	Number of	Number of
	ITT	Std. Err	Lower	Upper	Mean	Observations	Pods
A. Juvenile Detention Outcomes							
Not Released	-0.03	0.03	-0.09	0.03	0.10	253	30
Days to Release	0.02	8.19	-16.04	16.07	34.69	231	30
Any New Spells	-0.06	0.07	-0.19	0.07	0.58	253	30
Number New Spells	-0.207	0.13	-0.46	0.04	1.07	253	30
Standardized Index	-0.11	0.09	-0.28	0.07	0.01	253	30
B. Arrest Outcomes							
Total	-36.13	28.37	-91.74	19.49	206.29	253	30
Violent	-14.617*	8.06	-30.41	1.18	34.29	253	30
Property	-6.23	6.61	-19.19	6.74	20.57	253	30
Drugs	-7.386*	3.85	-14.94	0.16	18.29	253	30
Other	-7.90	21.00	-49.05	33.26	133.14	253	30
Standardized Index	-0.129*	0.07	-0.27	0.01	0.00	253	30
C. Social Costs of Crime							
Direct Cost, All Crimes	-43350	32554	-107155	20455	118228	253	30
Direct Cost, Arrests	-9911*	5510	-20710	888	24715	253	30
Willingness to Pay, All Crimes	-360676**	118673	-593275	-128077	567761	253	30
Willingness to Pay, Arrests	-45525**	16175	-77228	-13822	75379	253	30
D. Education Outcomes							
Attended any Days	0.07	0.05	-0.03	0.18	0.51	244	28
Number of Days	5.72	4.01	-2.13	13.57	21.60	244	28
Has GPA	0.136**	0.04	0.06	0.22	0.39	244	28
GPA	0.30	0.19	-0.07	0.66	1.10	105	28
Standardized Index	0.208**	0.06	0.08	0.33	-0.07	244	28

Notes: This table shows estimates of the intent-to-treat effect of the program for the second year of the program when random assignment is correlated with program participation. The second year includes 30 pod-lottery clusters. Standard errors clustered on individual and residential pod using the method described in Cameron and Miller (2015). All regressions are based on Equation 3.2 using all 626 youth in the study. The year one observations improve the precision of the estimated coefficients on baseline covariates but do not otherwise contribute to the treatment effect estimates. Stars indicate: * $p<0.1$, ** $p<0.05$, *** $p<0.01$.

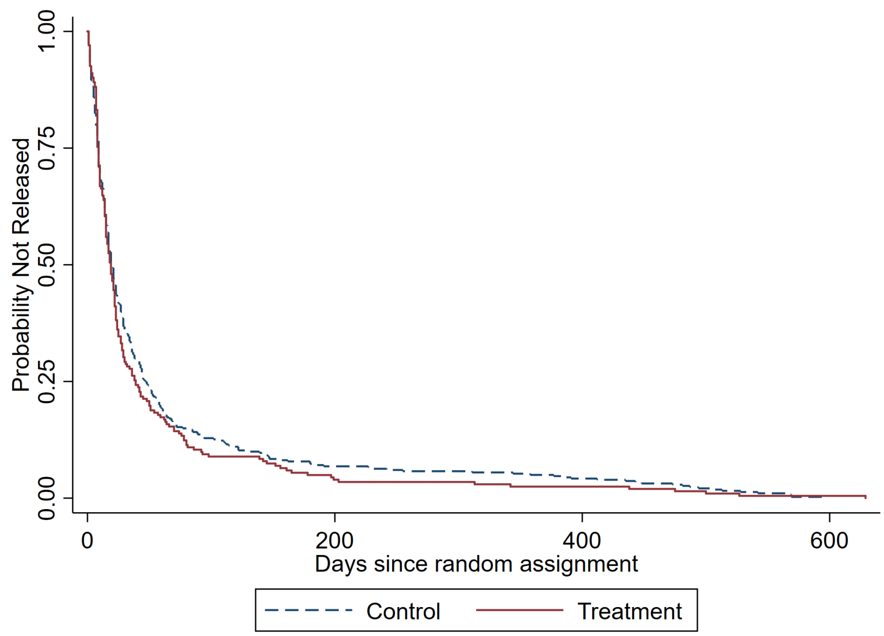


FIGURE 7. Empirical Survival Curve for Time Until Release

Notes. This figure shows the proportion of youth in the treatment and control group who had not been released after X days.

The first two rows of Table 8, Panel A provide additional evidence that there is no impact on the time to release. The first row shows that the youth in the treatment group were 3 pp less likely to have been released by the end of our court records data. Although this is a 30 percent reduction relative to the control group's mean of 10 pp, it is not statistically significant at conventional levels. The 95 percent confidence interval includes a reduction as large as 9 pp up to an increase of 3 pp. The second row shows the impact on days until release among the sub-sample which was released during our sample period. Treatment group youth were released 0.02 days later than the youth in the control group. This effect is also statistically insignificant with a 95 percent confidence interval ranging from a 16-day decrease up to a 16-day increase. The control group's average days to release is 34.69.

The third and forth rows of Panel A show the impacts on having any new spells in the JTDC in the first eight post-randomization months[14] and on the number of new spells during this period. Both point estimates are negative. The impact on having any new admissions to the JTDC in the first eight post-randomization months is not statistically significant, but suggests treatment caused a 6 pp reduction in the probability of having any new spells. This is a 10 percent reduction relative to the control group mean of 0.58. The confidence interval ranges from a 33 percent reduction to a 12 percent increase. The estimated impact on the number of post-randomization spells, -0.21, indicates the program caused a 19 percent reduction in the number of new spells. This effect is more precisely estimated than the other estimated impacts on juvenile incarceration ($p = 0.11$). The 95 percent confidence interval includes up to a 0.46 decline to a 0.04 increase in the number of new spells. Therefore, we can rule out that the program caused more than a 4 percent increase in the number of new spells, but it potentially caused as much as a 43 percent reduction in the number of new spells. Our point estimate is somewhat larger, but of a similar magnitude, to comparable intent-to-treat estimates on having any readmissions to the JTDC and on the number of readmissions to the JTDC after 8 months in Heller et al. (2017).[15]

The final row of Panel A reports the ITT effect on a standardized index of detention outcomes. This index averages standardized versions of the previous four outcomes. Being randomly assigned to treatment causes a 0.11 standard deviation (SD) reduction in the detention index. Because greater values for all of

[14]We use 8 months because this is how many months we observe new spells after the final randomization.

[15]They estimate that the reforms they study reduced the probability of any readmissions by 0.05 pp and the number of readmissions by 0.12 in the 8 months after random assignment (shown in their appendix).

the outcomes indicate more involvement with the justice system, a reduction is a beneficial outcome. The impact on this index is also statistically insignificant with the 95 percent confidence interval around the estimate ranging from a decline as large as 0.28 SDs up to an increase of 0.07 SDs.

Impact on Arrests

Readmission to the JTDC is an imperfect measure of youth's ongoing criminal behavior and interaction with the justice system. Youth could be arrested and released or if they have turned 18, they could be admitted to the adult jail. To better capture the extent of youth's ongoing criminal activity and involvement with the justice system Table 8, Panel B shows the impact of the program on total, violent, property, drug, and other arrests in the first year after the lottery. We look separately at these types of crime because they have different social costs and causes. It is not uncommon for programs to have different impacts on different types of crime (Davis and Heller, 2020). All outcomes are multiplied by 100 and are therefore interpretable as being the impact per 100 treatment youth. We also report impacts on the social cost of crime in Table 8, Panel C.

We estimate that the program caused a reduction of 36 arrests per 100 youth in the treatment group in the year after the program. This is an 18 percent reduction relative to the control group mean of 206 arrests per 100 youth. While large, this estimate is not significantly different from zero at conventional levels ($p = 0.20$). The 95 percent confidence interval around this point estimate ranges from a decline of 92 arrests per 100 youth to an increase of 19 arrests per 100 youth. So, consistent with the impact on future spells in the JTDC, we can rule out that the program caused more than a 9 percent increase in total arrests, but it potentially caused as much as a 44 percent reduction in total arrests.[16]

Figure 8 shows the dynamics of how the cumulative impact on total arrests in Table 8 develops in the first year after random assignment. The cumulative impact on total arrests is negative in each month and becomes more negative until 10 months after random assignment when it peaks at a reduction of nearly 50 arrests per 100 youth. Although the cumulative impact on total arrests is never statistically significant at the 5 percent level, the confidence intervals include large reductions and reject anything but small increases in total arrests.

Looking separately across types of crime, all of the point estimates in Panel B are negative. This uniformity in the sign of the effects is itself suggestive that the program reduces crime. Under the null hypothesis that the treatment effects are all zero with the noise distribution symmetric around zero, the probability of finding negative estimates for each type of crime is 0.06 ($= (1/2)^4$). With the exception of other crime, the effects are also quite large.

We estimate that the program caused a reduction of 15 arrests for violent crime per 100 youth in the year after the program. This is a 43 percent reduction relative to the control group mean of 34 violent crime arrests per 100 youth. This estimate is statistically significant at the 10 percent level ($p = 0.07$). The confidence interval around the point estimate ranges from a decline of 30.4 to an increase of 1.2 violent crime arrests per 100 youth. As with total arrests, the

[16] Appendix Tables A.4 and A.5 show the estimates pooling both years and in the first year of the experiment, respectively. Once again, we show these results for completeness and transparency, but do not expect to find any results in the first year given that treatment is not a relevant predictor of program participation that year. Not surprisingly, the first year impacts are all proportionally small and statistically insignificant. The largest effect in the first year suggests a 10 percent change.

confidence interval includes proportionally large reductions in violent crime (of up to 87 percent) and rejects all but a small increase (up to 3 percent).

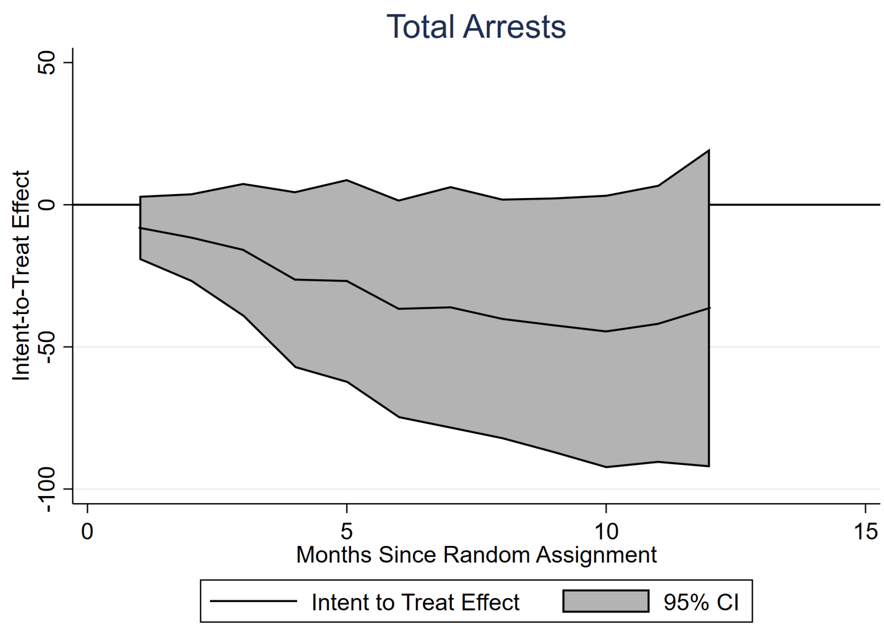


FIGURE 8. Dynamic Impacts of Intervention on Total Arrests

Notes. This figure shows the intent-to-treat estimates on number of total crime arrests over the first 12 months after random assignment for the second year of the study. 95% confidence interval based on standard errors clustered by pod and youth.

We find similarly large proportional reductions in arrests for property and drug crime. Our point estimates imply reductions of 30 and 40 percent, respectively. Specifically, the estimates suggest that the program caused 6.2 fewer property arrests and 7.4 fewer drug arrests per 100 youth. The drug crime estimate is significant at the 10 percent level ($p = 0.06$). The 95 percent confidence interval ranges from a decline of 15 drugs arrests to an increase of 0.2 drug arrests per 100 youth. Therefore, we can reject anything but a tiny adverse impact on drug arrests. The property crime estimate, however, is quite noisy with a 95 percent confidence interval ranging from a reduction of 19 arrests to an increase of 7 arrests per 100 youth. While a small increase in arrests for property and drug crime is not ruled out, the confidence intervals indicate that there may also be a large reduction.

The estimated impact on other crime is proportionally small, negative, and statistically insignificant. We estimate the program caused a reduction of 8 arrests per 100 youth in the treatment group in the year after the program. This is a 6 percent reduction relative to the control group mean of 133 arrests per 100 youth. The estimate is also noisy, with a 95 percent confidence interval ranging from a reduction of 49 arrests to an increase of 33 arrests per 100 youth.

Figure 9 shows the dynamics of how the cumulative impacts on arrests for violent, property, drug, and other crime shown in Panel B of Table 8 develop over the first 12 months after random assignment. The estimated impact on violent arrests begins negative and becomes more negative over the year. The estimate is statistically significant at the 5 percent level in each of the first 5 months and 9 months after random assignment. In all months, we can reject anything but a small increase in violent arrests. The impact on property arrests is positive until about 6 months after random assignment when it becomes negative. The impact on drug arrests starts negative and becomes more negative in the first 4 months after random assignment when it begins to level out or grow more negative at a slower rate. The drug impact is statistically significant at the 5 percent level in all but the 7th and 12th month after random assignment. Finally, the impact on other arrests starts at zero and becomes more negative until 5 months after randomization when it begins to gradually shrink towards zero.

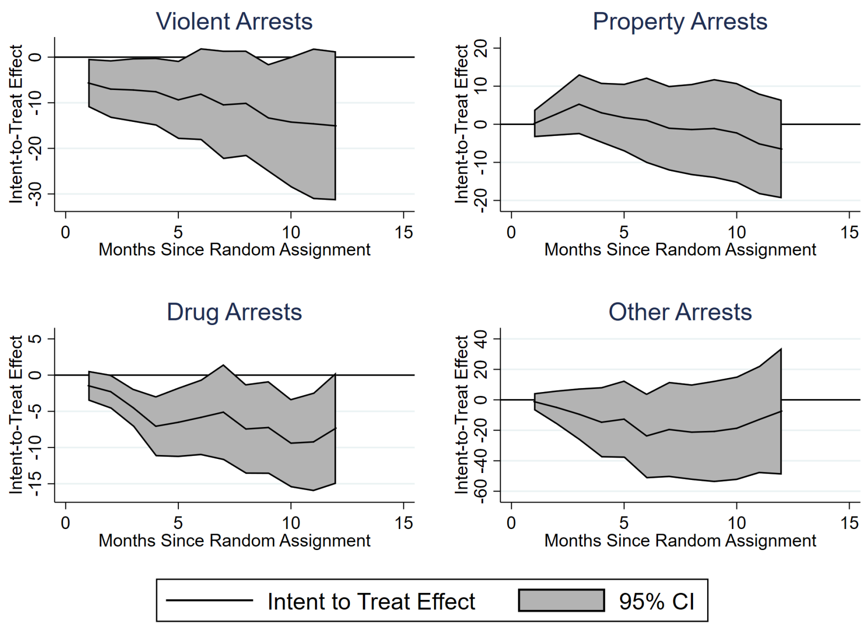


FIGURE 9. Dynamic Impacts of Intervention on Arrests by Type of Crime

Notes. This figure shows the intent-to-treat estimates on crime outcomes over the first 12 months after random assignment for the second year of the study. 95% confidence interval based on standard errors clustered by pod and youth.

The final row of Panel B shows the impact on our standardized index of arrest outcomes. This index averages standardized arrests for violent, property, drug, and other crimes. We find a marginally statistically significant 0.13 SD reduction in arrest outcomes ($p = 0.07$). The 95 percent confidence interval includes reductions as large as 0.27 SDs up to basically no effect, 0.01 SDs.

This standardized index puts equal weight on violent, property, drug, and other crimes. However, the social costs of these crimes are quite different. An alternative to this approach is to look at the program's impact on the social costs of crime directly. The social cost of crime outcome can be thought of as a weighted index of total arrests with greater weights given to more socially costly crimes. Panel C of Table 8 shows these results.

Assigning social costs to crime is an inherently noisy endeavor that relies on strong assumptions. We show the results for the same social cost of crime outcomes as are shown in the appendix of Davis and Heller (2020).[17] The first and second rows of Panel C use estimates of the direct costs of crime from Miller et al. (1996) as reported in Cohen and Piquero (2009). These are based on direct victim costs and direct costs to the justice system. The third and fourth rows of Panel C show results using contingent valuation based estimates of people's willingness to pay to avoid crime from Cohen et al. (2004) which are also reported in Cohen and Piquero (2009). To account for the fact that we observe arrests, not all crimes committed, the first and third rows inflate the number of arrests by the estimated ratio of crimes to arrests reported in Cohen and Piquero (2009). The second and fourth rows show estimates using only observed arrests.

Our estimates in Panel C are all large and negative, indicating reductions in social costs ranging from \$9,911 to \$360,676 per treated youth. The estimates indicate about a 40 percent reduction in the social cost of crime using the direct estimates and just over a 60 percent decline using the willingness-to-pay estimates. The willingness-to-pay estimates are both statistically significant at the 5 percent level. The least beneficial impacts in the 95 percent confidence intervals around these estimates suggest a roughly 20 percent reduction in the social costs of crime. The direct cost of crime estimate is statistically significant at the 10 percent level when using observed arrests, but not when using all crimes.

Scaling the point estimates by the 78 youth who were in the treatment group in the second year of the study suggests that the program has a potentially high rate of return. Even the smallest point estimate, \$9,911, suggests a total societal

[17] Costs are converted to 2012 dollars using the consumer price index.

savings of over $750,000 from just the second year of the forums. This would be a 50 percent return on the initial $500,000 grant to support PSN work in the Northern District of Illinois. Of course, the confidence interval around the total societal savings using the direct estimates includes no benefit and even some smaller detrimental effects. The magnitude of these benefits is plausible because crime is incredibly socially costly. A single armed robbery is estimated to cost the criminal justice system about $16,000 and the victim about $32,000. Cohen and Piquero (2009) report estimates suggesting that youth are arrested for about 1 out of every 2.8 armed robberies they commit. Therefore, a reduction of a single arrest for armed robbery is estimated to create societal savings of over $130,000 using the more conservative direct estimates.

Impact on Schooling

Table 8, Panel D shows the program's impact on post-randomization schooling outcomes among the 96 percent of youth who are linked to a Chicago Public School record in the second year of the experiment. The first and second rows of Panel D show the impact on attending any days at a non-prison public school in Chicago and on the number of days attended during the school year following random assignment, respectively.[18] We focus on non-prison schools because youth are required to attend Nancy B. Jefferson when they are incarcerated at the JTDC.

Youth in the treatment group are 7 pp more likely to have attended any days of school. This a 15 percent increase relative to the control group mean of

[18]For lottery 5, in April 2016, attendance is based on the 2015-16 school year. For lotteries 6 and 7 in June and August 2016, attendance is based on the 2016-17 school year.

51 percent but is not significantly different from zero. The 95 percent confidence interval ranges from a decline of 3 pp to an increase of 18 pp.

We also find that the program did not significantly increase the number of days attended. The point estimate is relatively small in absolute terms, an increase of 5.72 days of school attended or just over one extra week of school. But this is a 26 percent increase relative to the control group's very low average level of attendance of 21.60 days attended. The 95 percent confidence interval around the point estimate ranges from a decline of 2.14 days attended to an increase of 13.58 days attended. Mirroring the crime results, the confidence interval includes proportionally large increases in attendance of up to 63 percent and rejects all but relatively small decreases of up to 10 percent.

The third row of Panel D shows the program caused a 13.6 pp increase in the probability that youth have a GPA in the semester following random assignment. Having a GPA is a measure of school attachment because youth with very low attendance are disenrolled from school and not graded. This is a statistically significant 35 percent increase relative to the control group mean of 39 percent. The 95 percent confidence interval around the point estimate ranges from an increase of 6 pp up to 22 pp.

The fact that we observe GPAs more often in the treatment group than the control group complicates our estimates of the impact of treatment on GPAs. In particular, we may expect that the program pushed more marginal students into school who otherwise would not have attended enough to be graded at all. This would bias our estimate of the impact on GPAs downward. Even so, we find a proportionally large, but statistically insignificant, 0.3 grade point increase in GPAs. This is a 27 percent increase relative to the control group's mean of 1.1.

In grade terms, treatment shifts students from a D average to a D+ average. The confidence interval ranges from a decline of -0.07 grade points up to an increase of 0.66 grade points. The lower bound of this grade point is not large enough to correspond to a change in students' average grades. The upper bound, in contrast, corresponds to an increase from a D to a C- average.

The final row of Panel D shows the impact of treatment on a standardized schooling outcome. This outcome includes the average standardized value of each observed schooling outcome. In particular, GPA is included when it is observed but is otherwise excluded. This assumes GPAs are missing at random so that students who are missing a GPA would have performed about as well as students with observed GPAs, on average. This should be conservative given that we find evidence that treatment reduces missingness and improves performance. The estimate suggests treatment improves schooling outcomes by 0.21 SDs. This effect is statistically significant at the 5 percent level. The 95 percent confidence interval includes improvements of between 0.08 and 0.33 SDs. This demonstrates that, when taken together, these schooling results provide relatively strong evidence that the youth outreach forums improve attachment to schooling.

3.6 Discussion

This paper presents the results of a randomized controlled trial designed to credibly measure the impact of a series of "youth outreach forums." These forums give youth the opportunity to talk to members of their community, authority figures in their life, and individuals who have left a life crime in a neutral, constructive setting.

Our estimates suggest that being randomly assigned to the treatment group offered the youth outreach forums caused an 18 percent reduction in total arrests driven by 43 percent, 30 percent, and 40 percent reductions in violent, property, and drug crime, respectively. Although the violent crime and drug crime impacts are both statistically significant at the 10 percent level, the confidence intervals around all of these estimates reject all but small increases in arrests and include even larger proportional reductions. We find a marginally significant reduction in a standardized index of arrest outcomes that weights different crimes equally. Looking at the social costs of crime, which can be thought of as another arrest index that gives greater weight to more socially costly crimes, we find proportionally large and very socially valuable reductions in the social costs of crime. Three of our four social cost estimates are significant at the 10 percent level and two are significant at the 5 percent level. Our estimates also suggest that the forums may increase attachment to school.

The forums were motivated by the theory that giving youth the opportunity to discuss the rationale for the law with authority figures in their lives will increase their perception of its legitimacy (Tyler, 2006a; Bottoms and Tankebe, 2012). In practice, however, the forums can be viewed as part of a more general focused deterrence strategy. The program "pulls many levers" that could drive our results (Braga and Kennedy, 2021). One lever is improving youth's perceptions of legitimacy. Unfortunately, we were not allowed to survey the youth in our study about their attitudes and beliefs so we cannot directly test this. Nevertheless, several components of the program could operate through this channel. Meeting with police command staff provides youth with the opportunity to ask questions about the rationale behind the law and policing and also the rules that police are expected to follow themselves. Discussions with community members, parents, and teachers also provide insight to the logic behind laws and rules and promote the sense that these laws and rules are based on a shared set of values.

Another lever is deterrence (Becker, 1968; Nagin, 2013). Deterrence theory suggests individuals will commit a crime when the perceived benefits outweigh the perceived costs. Meeting with former offenders who spent decades of their lives in prison or who were paralyzed because of gun violence could both increase the perceived costs of committing crime. Meeting with command staff from the Chicago Police Department also highlights that the youth are under scrutiny which may suggest that they are more likely to be caught if they do commit a crime.

Yet another lever is connecting youth to social supports and building their sense of community belonging. Youth were connected with a school reengagement specialist working at the JTDC whose job was to help them transition back to the school that best meets their needs. But they also were introduced to other social programs operating in Chicago in the "Community and a Shared Moral Voice" forum. Being offered help shows youth that their community would like them to succeed and potentially addresses root causes of criminal behavior, like needing to make extra money to help pay bills. Moreover, almost all of the forums created opportunities for youth to meet positive adult mentors.

Finally, the program could also build skills that help the youth stay out of trouble and succeed in school. The "Hope Outside the Law" forum was replaced by a forum that explicitly taught the youth conflict resolution tools based on cognitive behavioral therapy principles. Indeed, credible experimental evidence has shown that teaching these types of skills can reduce crime (Landenberger and Lipsey, 2005; Heller et al., 2017).

The challenges that arose in this project are relevant for future researchers who hope to conduct similar experiments in juvenile detention centers. The first year of the study was uninformative about treatment effects because of short spells in the JTDC. By working with the JTDC's administration to run the randomization closer to the start of the program, we were able to correct this issue in the second year.

Our estimates are somewhat imprecise because of the issues that arose in the first year. Given the imprecision, we take care to emphasize effect sizes and confidence intervals in addition to p-values. Recent debates in the statistics literature on the value of hypothesis testing have emphasized that a focus on effect sizes and confidence intervals may better address whether a hypothesis is correct than the traditional focus on p-values (Wasserstein and Lazar, 2016; Benjamin et al., 2018). Therefore, this may be a useful template for other criminology research.

CHAPTER IV

DO BUDGETS REFLECT PRIORITIES? THE IMPACT OF SCHOOL FUNDING REFORM ON DISTRICTS AND TEACHERS IN WASHINGTON STATE

4.1 Introduction

Education spending has historically been one of the largest components of state and local government budgets. It was the largest portion of state and local expenditures from 1977 to 2014, before being surpassed by public welfare spending (Urban Institute, 2021). As of 2021, however, kindergarten through 12th-grade education expenditures still made up a large portion of state spending at 20.5% (Urban Institute, 2021). This substantial allocation of both state and local funds—combined with frequent changes in policy affecting how school districts are funded—has led many researchers to explore topics in this area.

School districts face pressure from government entities and the public to perform well and allocate resources strategically (Grissom et al., 2017; Booher-Jennings, 2005; Clotfelter et al., 2006; Dee et al., 2013). Many studies, including older observational work and more recent causal analyses, show that increased education spending improves outcomes for students in both the short and long run (Jackson, 2018).

This paper explores the impacts of a school finance reform (SFR) in Washington state. In 2018, the funding formula for how school districts were funded by the state changed, leading many school districts to increase spending. The funding formula changed such that each district was exposed to the SFR differently. I examine how school districts spent their additional funds after the SFR. A heterogeneity analysis is conducted to examine differences in spending across districts. I then examine the funding change and if it is predictable on a set of characteristics. Lastly, I examine whether teacher retention is affected by a school district's exposure to the policy change.

I begin with an accounting exercise to assess which expenditure categories increased when school districts raise their overall spending. I show that after the SFR, districts do not change how they allocated the marginal dollar. I then calculate exposure to the SFR by finding the difference between the estimated apportionment under 2017-2018 characteristics in the new funding formula and the actual apportionment for the 2017-2018 school year. I use this measures to proxy for exposure to the SFR to calculate if differences in exposure to the SFR increase spending in certain categories. I find that each additional dollar of exposure is generally associated with increased school spending on certificated staff salaries. I also conduct heterogeneity analyses to show that school districts with an exposure above the median tend to spend less on certificated salaries with their additional exposure dollars. Further, districts with a positive locality multiplier tend to allocate more of their additional apportionment to certificated staff salaries compared to districts without such a multiplier.

Evaluating how school districts respond to a SFR or some other financial shock has been a topic of interest to many researchers. Jackson et al. (2016) find that positive school district funding shocks disproportionately increase instructional spending. To contrast, Jackson et al. (2021) find that when there is a negative spending shock, districts distribute the decrease in funding similarly across capital expenditures and operational costs. There is also evidence that richer and poorer districts spend additional funds in different ways. For example, high-poverty schools were more likely to use ESSER funds for facility improvements, whereas low-poverty schools directed more of their funds toward mitigating learning loss (Brooks and Springer, 2024). However, Springer et al. (2009) find that there is no statistical difference on how school districts spend their money after a court equity vs. adequacy reforms. Card and Payne (2002) find that increasing aid to poorer districts decreases the spending gap between high and low income districts.

Most research involving SFRs focuses on a second order outcome (student outcomes, teacher outcomes, etc.), and focuses on one category of expenditure of SFRs. Understanding these responses is crucial, as SFRs and similar shocks have been linked to student outcomes. For instance, Jackson and Mackevicius (2021) conduct a meta-analysis of 31 studies looking at student outcomes from a school spending increase and find that 90 percent of the time if the funding increase is persistent (4 or more years), then the funding increase increased educational attainment or test scores. While this study does not directly examine student outcomes, it can be put in the context for other results which find that targeted spending increases increase different student outcome measures (Lafortune et al., 2018; Candelaria and Shores, 2019; Kreisman and Steinberg, 2019).

This paper also examines how increased school district apportionment allocations impact of teacher retention. I show that within-district teacher retention is associated with increases in estimated apportionment after the SFR. My results are consistent with those of Sun et al. (2024) who find short-run funding-induced increases of teacher retention after the McCleary Decision. Zamarro et al. (2024) study a direct teacher salary increase and find minimal positive effects on retention after the first year. While the mechanism isn't a SFR in Hendricks (2014), the study does establish a causal effect between teacher pay and teacher retention.

The remainder of the paper is organized as follows. Section II provides a brief history of the McCleary Decision. Section III explains each of the data sources used. Section IV discusses the methodology used in my analysis. Section V provides results. Section VI concludes.

4.2 Background

Education in the United States is funded from three primary sources: local, state, and federal. The majority of education funding comes from state and local sources. At the state level, there are two primary models that states use to fund education: student-based and resource-based models. In student-based funding models, the state provides a base amount of funding per student. In resource-based models, states institute standards to which they will fund different aspects of basic education. For example, in Washington state, teachers for kindergarten through 3rd grade are funded by the state at 1.2 FTE per 17 students. The state then has a base salary amount per district that they use to compute total certified instructional staff funding. Other resources such as staff and technology are funded in a similar way.

Washington's teacher funding mechanism changed significantly as a result of the McCleary Decision, a landmark ruling by the Washington State Supreme Court. McCleary v. Washington was filed in 2007 by two Washington families, who alleged that the state was not fulfilling its constitutional obligation to fully fund basic education.

In January 2012, the Washington State Supreme Court ruled in favor of the plaintiffs and ordered the State of Washington to fully fund education. This order became known as the McCleary Decision. While the court did not require an immediate remedy, they did require the state to make a plan as to how they would comply with the Washington State Constitution. In the fall of 2014, the Washington State Supreme Court found the state in contempt of court for failing to comply with the court order. Additional funds from the state were provided to districts starting in the 2017-18 school year, with all changes being fully implemented by the 2019-2020 school year, one year later than the State Supreme Court had mandated.

The legislature's response came in the form of Engrossed House Bill 2242, which increased state funding to school districts. Each district receives a different amount of money from the state based on many factors, and this money is divided into three spending categories: General Purpose, Special Purpose, and Capital Grants.[1]

One of the most significant changes enacted by the bill was a new formula for allocating teacher salaries to districts in 2018. Teacher salaries are funded through general purpose apportionment funds. All funds that a school district receives through apportionment are general purpose, and apportionment consists of three categories: funding for employee salaries, MSOC (materials, supplies and operating costs), and remote/small school/district adjustments. In 2018, Washington transitioned from a staff mix model to a locality-based model for allocating teacher salaries. The staff mix model finds an allocation based on school district enrollment and then takes into account every teacher employed by the

[1]General Purpose funds are allocated with recommendations from the state but are ultimately at the discretion of the school district. Special Purpose funds are restricted to specific uses.

district's experience and highest degree earned to determine a salary allocation for that teacher. For the 2018–2019 school year, each district received $65,216 per teacher FTE, multiplied by its regionalization factor. This factor—ranging from 0.00 to 0.24—is based on residential property values in the district. Districts with housing costs above the state median received a factor of 0.06, 0.12, 0.18, or 0.24, while those below the median received 0.00.[2] Beginning in the 2019-2020 school year, an additional experience factor of 0.04 experience factor was added to allocations of teacher salaries for districts whose average teacher experience and education exceeded the state average. A full breakdown of the 2018–2019 apportionment formula is provided in Section 4.6.

Researchers have begun studying the effects of this funding formula change on various outcomes. This paper most closely relates to Sun et al. (2024), as they leverage the funding formula change to calculate a measure of exposure, and they use that exposure mechanism to calculate it's effect on teacher salaries, hiring and retention. I focus on the first order question of examining how exposure to the policy impacts school district expenditures, while their paper examines the more-specific, second order question of how teachers were affected by the policy specifically examining heterogeneous effects for how teachers' pay was impacted, the change in teacher turnover rates and the change in teacher hiring. This paper and Sun et al. (2024) also use different methods to calculate exposure to the SFR. This paper calculates the exposure to the SFR as the difference between the amount that school districts receive under the new and old funding formulas based on district characteristics in the 2017-2018 school year. This is a continuous measure of exposure. Sun et al. (2024) use a discrete measure of exposure by

[2]The regionalization factor applies to all certificated employee salaries, not just teacher salaries.

breaking up exposure into terciles. Sun et al. (2024) also define exposure to the SFR differently. They calculate the estimated staff apportionment for staff using estimated 2018-2019 district enrollments, then divide by the number of students to get a "dosage" amount per student. They group districts into terciles based on their dosage amount. This method does not account for the old funding formula.

4.3 Data

School district finance data were obtained from the Washington State Office of the Superintendent of Public Instruction (OSPI) and the Washington State Fiscal Information website. OSPI collects these data annually through required district reports. All dollar values are adjusted for inflation using the CPI-U and expressed in 2024 dollars.

District-level expenditure data are available from the 1990–1991 to 2023–2024 school years. Revenue data are also available for the same years. One district appears to have a data reporting outlier in the 2021–2022 school year and is therefore excluded from the main analysis.

Panel A of Table 9 presents summary statistics of expenditures by object for the school years 1990-1991 to 2023-2024. The "Exp./Pupil" column shows the average per pupil spending for each category. The "Share Total Exp." column shows the average share of the total spending of that category. The "Δ Exp./Pupil" shows the year-over-year difference of expenditure in each category. On average, the largest category of expenditure is salaries for certificated staff, which are more than double the amount spent on any other category. Despite some noise, the average year-over-year change in per-pupil expenditure is positive across most categories, indicating increased investment by school districts. In Panel B of

Table 9 provides similar statistics for district revenues by major funding source. These data can be obtained on the Washington State Fiscal Information website. A large percentage of school district revenue, about 68%, comes from state sources. Note that the share statistic does not add up to 1 due to the omission of small categories and rounding.

TABLE 9. Expenditure and Revenue Summary Statistics

Panel A: Expenditure Object	Count	Exp./Pupil	Share Total Exp.	Δ Exp./Pupil
Salaries - Certificated	10131	6951.38	0.42	82.96
		(2961.38)	(0.06)	(1815.91)
Salaries - Classified	10131	2907.13	0.17	56.90
		(1656.31)	(0.03)	(835.24)
Employee Benefits and Payroll Taxes	10131	3338.08	0.19	73.85
		(1781.37)	(0.03)	(966.30)
Purchased Services	10131	2432.80	0.13	88.72
		(3013/25)	(0.06)	(2828.68)
Supplies	10131	1198.08	0.07	27.24
		(852.43)	(0.02)	(565.33)
Travel	10117	80.07	0.00	-0.11
		(116.67)	(0.00)	(88.14)
Capital Outlay	10109	224.81	0.01	-7.20
		(488.85)	(0.02)	(551.04)

Panel B: Revenue Sources	Count	Rev./Pupil	Share Total Rev.	Δ Rev. Pupil
State	10141	12898.65	0.74	185.56
		(7689.15)	(0.09)	(4944.01)
Federal	10141	1840.41	0.10	53.64
		(2335.26)	(0.08)	(969.52)
Local	10131	2474.81	0.15	35.86
		(1830.41)	(0.07)	(1362.40)

Notes: This table shows the summary statistics for expenditure categories (Panel A) and the 3 main sources of revenue for school districts (Panel B) from the 1990-1991 school year to the 2023-2024 school year. Change in expenditure per pupil is shown from the 1991-1992 school year to the 2023-2024 school year. All values are in 2024 dollars.

There are two ways which district expenditures are classified: by object category and by activity code. Object categories are more general. There is not a direct mapping of activity codes to object categories. For example, the activity code 27, which is for teaching, includes both pay to classroom teachers and supplies for both instruction and students. However, this code does not include textbooks as there is a separate curriculum code.

Additional district-level characteristics were obtained from the National Center For Education Statistics (NCES).

The main source of employment information in Washington state public schools is the S-275 aggregated personnel reports from the OSPI for the 2013-2014 school year to 2023-2024 school year.[3] S-275 is a form filed by each school district for each employee and then submitted to the OSPI. This includes employees in non-teaching roles. Key components of this form include a teacher's highest attained degree, years of professional experience, salary information, and detailed specifications of their work assignment.[4] This data is then compiled by the OSPI, and is made publicly available in Microsoft Access and Excel files.

I compile these reports for the 2013-2014 school year to the 2023-2024 school year. To focus on teacher retention from this SFR, I restrict the sample to teachers and employees who were a teacher at some point in the sample, but who currently have another role.

Teacher retention can be measured in several ways: within the same school, within the same district, within the teaching profession, or within the broader education profession including administrative roles. Because the McCleary Decision changed salary allocations from the state to the districts, the main retention indicator that I create is for teachers who stayed teaching in the district. In creating this variables, I drop the 2013-2014 school year, since I do not have the previous school year to indicate retention.

[3]2023-2024 data is preliminary.

[4]An example of the form can be found here: https://ospi.k12.wa.us/sites/default/files/2023-08/b079-21attach2.pdf

4.4 Methods

Marginal School Expenditures

The first model functions as an accounting exercise to assess whether school districts altered their use of the marginal per-student dollar following the McCleary Decision. Specifically, I estimate whether the composition of expenditures changed post-reform. The model includes school district and year fixed effects:

$$\Delta ObjExp_{d,t} = \gamma_1 Post * \Delta TotalExpenditure_{d,t} + \alpha_t + \beta_d + \epsilon_{d,t} \tag{4.1}$$

In Eq. (4.1) $\Delta ObjExp_{d,t}$ represents the change in object-specific expenditures per student for certificated salaries, classified salaries, employee benefits, supplies, services, travel and capital outlay from year $t-1$ to school year t in district d. $\Delta TotalExpenditure_{d,t}$ is the change in total expenditure per student of district d in school year t. $Post$ is an indicator for school years post SFR where $Post = 1$ for $t > 2017 - 2018$. α_t are school year fixed effects and β_d are school fixed effects.

SFR Funding Changes and Future Spending

To evaluate whether changes in district funding due to the SFR affected expenditure levels in subsequent years, I extend the model to include a district-level exposure measure and estimate:

$$Y_{d,t} = \gamma_1 Post * Exposure_d + \alpha_t + \beta_d + \epsilon_{d,t} \tag{4.2}$$

To better isolate this relationship of spending changes, I employ an instrumental variables (IV) approach using Post $\times$ Exposure$_d$ as an instrument for actual per-pupil expenditure.

First Stage:

$$Expenditure/Pupil_{d,t} = \gamma_1 Post * Exposure_d + \alpha_t + \beta_d + \epsilon_{d,t} \tag{4.3}$$

Second Stage:

$$Y_{d,t} = \theta_1 Post * \widehat{Expenditure}/Pupil + \alpha_t + \beta_d + \epsilon_{d,t} \tag{4.4}$$

In Eq. (4.2) and Eq. (4.4) $Y_{d,t}$ is the school year object of expenditure (certificated salaries, classified salaries, employee benefits, supplies, services, travel and capital outlay) spending per student in district d in school year t. In Eq. (4.2) and Eq. (4.3), $Exposure_d$ can be defined in various ways. I describe three possible definitions below. First, it could be the change in total expenditure per student of district d from the 2017-2018 school year to the 2018-2019 school year. Second, $Exposure_d$ could be the change in total apportionment per student of district d from the 2017-2018 school year to the 2018-2019 school year. Third, $Exposure_d$ is the difference in estimated apportionment using the new funding formula and district characteristics from the 2017-2018 school year and the actual apportionment for the 2017-2018 school year. $Post$ is an indicator for school years post SFR where $Post = 1$ for $t > 2017 - 2018$ and 0 otherwise, and α_t are school year fixed effects and β_d are school district fixed effects.

The first two exposure definitions reflect realized changes but may be confounded by non-random spending behavior or local factors. The third definition provides a more policy-pure measure by holding student and staff characteristics constant and applying only the new funding rules.

Because the largest shift in state funding was in teacher salary allocations, I focus on re-estimating apportionment related to teacher staffing. This allows for a better measure of exogenous funding variation attributable to the SFR.

Since how teachers were funded from the state drastically changed as a result of this SFR, I will use the elements in the teacher element of the school funding formula that changed to re-estimate the exposure variable. Using the data on revenue, I use apportionment revenue instead of expenditure and this was the revenue source impacted by the SFR. I calculate the new allocations for teachers in Eq. (C.1).

Though teacher salaries are the largest component, other apportionment categories (administrative, classified, and district-wide staff; MSOC; and adjustments for remote/small schools) are also critical. Funding levels for administrative and classified staff are \$95,000 and \$45,912 per FTE, respectively. MSOC is funded per pupil at \$1,264.07 (grades K–8), \$1,437.71 (grades 9–12), and \$1,495.56 for students in Career and Technical Education or Skill Center programs. After staff FTEs are calculated, a 5.3% adjustment is added for central district administration. The prototypical basic apportionment formula is shown in its simplified form in Eq. (C.12).

After estimating the apportionment under the new formula for the 2017-2018 school district characteristics ($\widehat{G(x)}$), I take the difference between that and the actual apportionment ($F(x)$) during the 2017-2018 school year to estimate the difference in apportionment under the two models, holding students fixed:

$$\widehat{G(x)} - F(x)$$

From there, I divide by the number of students in the 2017-2018 school year to get the change in per student apportionment under the two models in the 2017-2018 school year, which gives me a measure of exposure variable. I graph the relationship between the actual apportionment difference and the estimated apportionment difference in Fig. 10.

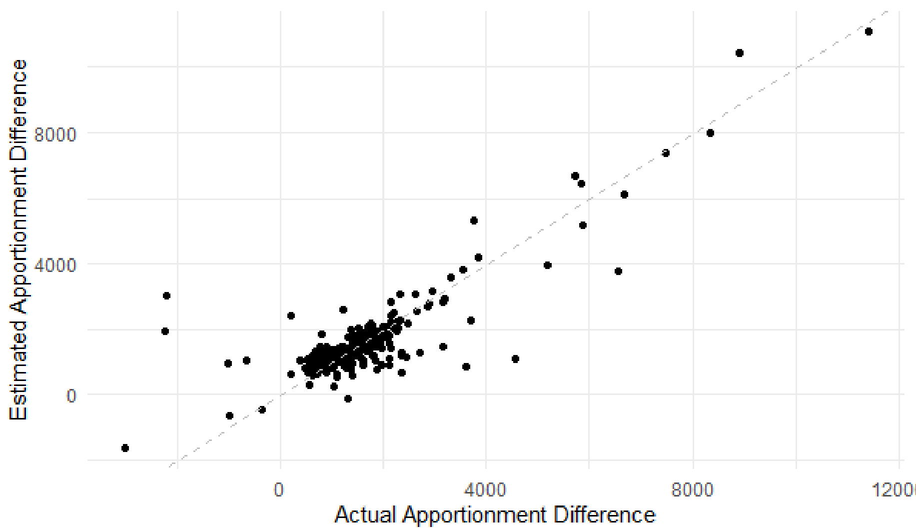


FIGURE 10. Apportionment Actual vs. Estimated Difference Per Pupil

Notes. This figure shows the actual difference in apportionment between the 2017-2018 and 2018-2019 school years on the x-axis and the estimated difference in apportionment between the estimated apportionment with the 2017-2018 school district characteristics under the new funding formula and the actual apportionment in the 2017-2018 school year.

Two key limitations affect the accuracy of estimated apportionment. One limitation is that I do not have student enrollment counts for CTE or skills centers. Therefore, estimated district apportionment may be underestimated specifically if there are a lot of students participating in these programs. A second limitation that also poses a risk for underestimation for apportionment is that I am not able to address small district or small school adjustments. To mitigate this concern, I add the difference between the actual and estimated apportionment for the 2018-2019 school year into $\widehat{G(x)} - F(x)$ which will account for small school and district adjustments assuming that this does not change from the 2017-2018 to 2018-2019 school years.

4.5 Results

Marginal Spending Increases

Table 10 presents the results from estimating the per pupil marginal change of each expenditure category on the per pupil marginal change of total expenditure, as specified in Eq. (4.1).

There are no significant changes in how school districts distribute their spending after the SFR. The coefficient on the change in per pupil total expenditure is statistically significant at the 1% level for all outcomes except that for capital outlay which is significant at the 5% level. When an indicator for the post SFR period is interacted with the change in per pupil total expenditure, none of the coefficients are statistically significant. This indicates that districts, on average, did not change the distribution of marginal spending between pre- and post-reform periods. On average, school districts spend \$0.35 of their marginal dollar on certificated salaries. The 95% confidence interval for this is approximately \$0.31 to \$0.38.

TABLE 10. Change in Per Pupil Spending

Panel A: Accounting	Δ Cert. Salary	Δ Clas. Salary	Δ Benefits	Δ Supplies	Δ Services	Δ Travel	Δ Cap. Outlay
Δ Per Pupil Total Expenditure	0.3448***	0.1535***	0.1825***	0.0771***	0.2183***	0.0048**	0.0192**
	(0.0119)	(0.0092)	(0.0137)	(0.0119)	(0.0053)	(0.0021)	(0.0086)
Post × Δ Per Pupil Total Expenditure	-0.0063	-0.0198	0.0008	-0.0115	0.0212	0.0033	0.0123
	(0.0234)	(0.0198)	(0.0161)	(0.0164)	(0.0233)	(0.0045)	(0.0138)
Observations	9,780	9,780	9,780	9,780	9,780	9,772	9,764
Panel B: TWFE	Cert. Salary	Clas. Salary	Benefits	Supplies	Services	Travel	Cap. Outlay
Post × Δ Est. Apportionment	0.4378***	0.1521**	0.2868***	-0.0041	0.2382***	-0.0012	-0.0340***
	(0.1680)	(0.0643)	(0.1003)	(0.0244)	(0.0899)	(0.0029)	(0.0127)
Observations	9,701	9,701	9,701	9,701	9,701	9,701	9,693
Panel C: IV	Cert. Salary	Clas. Salary	Benefits	Supplies	Services	Travel	Cap. Outlay
$Total\widehat{Expenditure}PerPupil$	0.4070***	0.1414***	0.2666***	-0.0038	0.2214**	-0.0011	-0.0316**
	(0.0768)	(0.0394)	(0.0419)	(0.0229)	(0.0946)	(0.0027)	(0.0151)
P-Value	0.38	0.78	0.12	0.02	0.8	0.02	0.02
Observations	9,701	9,701	9,701	9,701	9,701	9,701	9,693
Year FE	Y	Y	Y	Y	Y	Y	Y
District FE	Y	Y	Y	Y	Y	Y	Y

Notes: This table shows results from regressions of first-differenced real per pupil spending of different categories on the real overall per pupil spending difference, the interaction interaction of a post dummy and the real overall per pupil spending difference, year fixed effects, and school district fixed effects. These results drop one district due to an outlier. Standard errors are clustered at the district level. Data from Washington State OSPI. Stars indicate * p <0.1, ** p <0.05, *** p <0.01.

To further investigate the policy's effect on the distribution of marginal spending, I re-estimate the model using alternative definitions of the post-policy period. I begin by including all pre-policy years and limiting the post-policy period to the 2018–2019 and 2019–2020 school years. I then sequentially add subsequent years until the full post-policy sample is restored. Results are shown in Fig. 11. This figure shows that school districts didn't significantly change how they were spending the marginal dollar during the post SFR period.

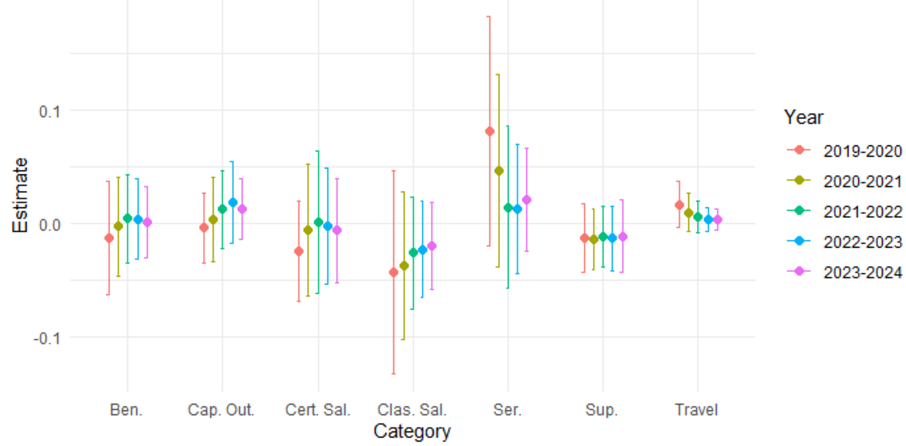


FIGURE 11. Regression Coefficients by Last Year Included

Notes. This figure shows the coefficients for the regression of the change in spending in a category on the total change in spending interacted with a post term. The year is the last year included in the sample. Error bars show the 95% confidence interval.

Exposure to SFR

To further investigate the effects of the SFR, I explore how differences in district-level exposure to the reform influenced spending patterns. The results presented in Table 10, Panels B and C. The results in Panel B are those from a two-way fixed effects regression of an amounts spent per category on indicator for school years after the SFR and the exposure to the SFR calculated by the estimated change in apportionment. The results in Panel B are descriptive and illustrate associations between exposure and spending, but they do not account for potential endogeneity in district-level responses to funding changes. To address this limitation and better identify causal effects, I implement an instrumental variables approach in Panel C, using interactions between post-reform years and pre-determined exposure measures to instrument for total expenditure. Results from this analysis provide evidence that increased exposure to the SFR led to greater spending on certificated salaries, consistent with the policy's intended goal of boosting teacher compensation. While defining exposure as the difference between the estimated apportionment under the new formula and the actual apportionment in the 2017-2018 school year is my preferred specification, I show results for alternative definitions of exposure in Table A.6 and Table A.7.

All three exposure definitions yield consistent results: greater exposure to the SFR is associated with increased spending on certificated salaries. These findings support prior literature suggesting that increased funding from SFRs leads to greater investment in teacher compensation (Jackson et al., 2016).

In examining other outcomes, most show a consistent positive increase in spending on that category in the years following the SFR when spending or apportionment increases. Panel B and Panel C show an increase in benefit spending, which is consistent with results of an increase of spending on certificated salaries and classified salaries, since some benefits are tied to salaries.

While there are positive, significant results to the amount of spending on certificated salaries, classified salaries, benefits, and services, expenditure on capital outlay decreases as school districts have larger estimated changes in apportionment due to the SFR in the post-SFR period. While the estimates are relatively small in magnitude, -0.034 and -0.032 respectively, the sign change is interesting, especially compared to the coefficient in the accounting exercise. In the post-SFR period for the accounting exercise (Panel A), a 1 dollar change in school expenditure is associated with an approximately \$0.03 increase in spending on capital outlay.

To better understand the differences in the distribution of spending between Panels A and C, I calculate a bootstrapped p-value for the difference between the estimate in Panel C and the estimate in Panel A. This allows me to compare if the marginal propensity to spend induced by the policy (Panel C) is the same as the average marginal expenditure (Panel A). P-values are calculated with a bootstrapped sample with replacement and 1,000 iterations. Results are shown in Panel C. I find no evidence that the estimates in Panel C compared to Panel A are statistically different for certificated salaries, classified salaries, benefits, and services. I do find evidence that the estimates are different for supplies, travel, and capital outlay.

To further explore the heterogeneity of the funding changes, I re-estimate Eq. (4.3), and Eq. (4.4) using expenditure data that is coded by activity which yields more regressions as there are more activity codes than object classifications. Expenditures per activity code can fall under multiple object classifications, so there is not a 1-to-1 mapping. I show estimated coefficients and 95% confidence intervals in Fig. 12.

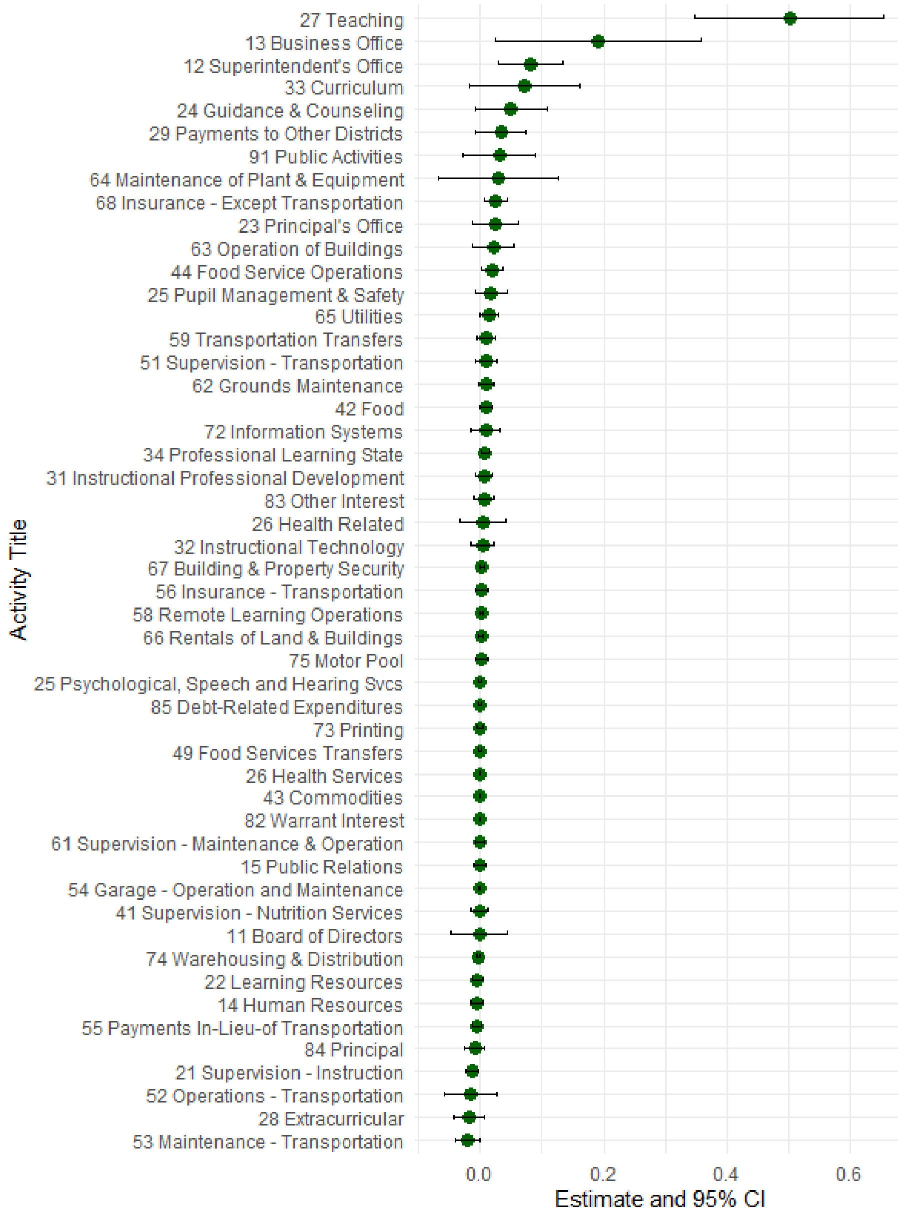


FIGURE 12. Activity Code Exposure IV Regression

Notes. This figure shows the coefficients for the IV regression where exposure to the SFR interacted with an indicator for post is used to instrument for change in total expenditure for the 2017-2018 to 2018-2019 school year. The second stage estimates the effect on the change in expenditure per activity code. Error bars show the 95% confidence interval.

These results start to examine more heterogeneity in where the changes in spending occurred for higher exposed districts. It is consistent with the results in Table 10 that these districts spent more funds on teaching and the superintendent's office as within those activity codes are salaries to teachers and school administrators. There are many activities that contribute to expenditures on services that had significant impacts in the activity code regression. These activity titles include business office, insurance, food service operations, and professional learning. There is also noisy evidence that expenditure in curriculum, guidance & counseling, payments to other districts, operations of buildings, and pupil management & safety increases with increased exposure to the SFR. All of these activity titles have a component that relates to purchased services.

Distributional Effects

I now explore the heterogeneity of the funding effects. First, I regress the estimated change in per-pupil apportionment on variables used in the old and new funding formulas. Table 11 shows that indicators for locality multipliers—key components of the new funding formula—are strong predictors of funding increases, consistent with policy design. Districts with higher multipliers received larger per-pupil increases.

To test whether high-exposure districts allocated funds differently, I split the sample at the median exposure value ($1,248.60) and estimate both the fixed effects model and the instrumental variables model.

TABLE 11. Characteristics Predictive of Apportionment Change

	Δ Per Pupil Apportionment
Total Students	-0.0501**
	(0.0193)
Staff Mix Factor	-2,228.0
	(1,385.8)
Locality Mult. 6%	692.9
	(425.5)
Locality Mult. 12%	781.6***
	(274.5)
Locality Mult. 18%	1,791.9***
	(527.3)
Locality Mult. 24%	1,492.6***
	(226.4)
Observations	302
R^2	0.14381

Notes: This table shows results a regression of the change in the estimate change in per pupil apportionment for the 2017-2018 school year to the 2018-2019 school year by district on different characteristics that are the primary factors in the school funding formula. Standard errors are clustered at the district level. Data from Washington State OSPI. Stars indicate * p <0.1, ** p <0.05, *** p <0.01.

Panel A of Table 12 presents results of regressions with different categories of spending as outcomes. The results suggest modest differences in how high-exposure districts allocate funds, although many coefficients are not statistically significant. However, in the IV specification (Table 13), districts above the median in estimated apportionment changes spend significantly more on certificated salaries, classified salaries, and benefits.

TABLE 12. Heterogeneous Treatment Effects by Apportionment Change and LEAP Multiplier

Panel A	Cert. Salary	Clas. Salary	Benefits	Supplies	Services	Travel	Cap. Outlay
Post × Δ Est. App.	0.8924***	0.5348***	0.4306***	0.0838	-0.4829***	-0.0048	0.0465*
	(0.2427)	(0.1752)	(0.1405)	(0.1010)	(0.0976)	(0.0070)	(0.0266)
Post × Above Median	246.8	444.4*	34.26	114.2	-663.6***	-5.683	103.8**
	(485.9)	(241.7)	(283.1)	(129.3)	(237.3)	(8.624)	(49.21)
Post × Δ Est. App. × Above Median	-0.4475	-0.4002**	-0.1373	-0.0931	0.7368***	0.0023	-0.0852***
	(0.3100)	(0.1887)	(0.1821)	(0.1046)	(0.1393)	(0.0074)	(0.0301)
Panel B							
Post × Δ Est. App. × Locality Mult. 6%	-0.3509	-0.0522	-0.3930***	0.0080	0.2213	0.0084	0.0315*
	(0.2305)	(0.1118)	(0.1146)	(0.0377)	(0.1583)	(0.0059)	(0.0187)
Post × Δ Est. App. × Locality Mult. 12%	0.6848**	0.4758**	0.3686***	0.0905*	-0.0560	0.0016	0.0620
	(0.3255)	(0.1863)	(0.1196)	(0.0505)	(0.2281)	(0.0094)	(0.0545)
Post × Δ Est. App. × Locality Mult. 18%	0.2004	0.0596	0.0154	-0.1077**	-0.2298	0.0026	0.0017
	(0.3950)	(0.1271)	(0.1798)	(0.0509)	(0.1766)	(0.0060)	(0.0248)
Post × Δ Est. App. × Locality Mult. 24%	1.373	-0.4971	-0.2327	0.0078	-0.3044	0.0183	0.1218*
	(1.152)	(0.3856)	(0.5947)	(0.1121)	(0.3705)	(0.0142)	(0.0704)
Year FE							
District FE							
Observations	9,701	9,701	9,701	9,701	9,701	9,701	9,693

Notes: This table shows results from a regression of real per pupil spending of different categories on the difference between the estimated apportionment in the 2017-2018 school year under the new funding formula and the actual apportionment in the 2017-2018 school year, a post-dummy, an indicator for if the estimated change in apportionment was above the median (Panel A) or an indicator for the locality multiplier used in the new funding formula (Panel B) and their interaction, year fixed effects and district fixed effects. Standard errors are clustered at the district level. Data from Washington State OSPI. Stars indicate * p <0.1, ** p <0.05, *** p <0.01.

TABLE 13. Heterogeneous Treatment Effects by Apportionment Change Amount and LEAP Multiplier: IV

	Cert. Salary	Clas. Salary	Benefits	Supplies	Services	Travel	Cap. Outlay
$\widehat{TotalExpenditure}$	0.5208	0.0701	0.3905**	-0.0490	0.3703*	-0.0050	-0.0818***
	(0.3696)	(0.1064)	(0.1774)	(0.0587)	(0.1962)	(0.0044)	(0.0227)
$\widehat{TotalExpenditure}\times$ Above Median	0.1170***	0.0537***	0.0530***	0.0052	0.0051	-0.0005	0.0072*
	(0.0306)	(0.0151)	(0.0182)	(0.0079)	(0.0255)	(0.0007)	(0.0043)
Observations	7,910	7,910	7,910	7,910	7,910	7,910	7,910
Panel B							
$\widehat{TotalExpenditure}$	0.2606	0.1432	0.3283**	0.0255	0.3238	-0.0030	-0.0408*
	(0.2470)	(0.1094)	(0.1427)	(0.0441)	(0.2043)	(0.0057)	(0.0207)
$\widehat{TotalExpenditure}\times$ Locality Mult. 6%	0.0360	-0.0494***	-0.0715***	-0.0284***	-0.1070**	-0.0012	0.0033
	(0.0298)	(0.0165)	(0.0196)	(0.0076)	(0.0454)	(0.0014)	(0.0040)
$\widehat{TotalExpenditure}\times$ Locality Mult. 12%	0.1638***	0.0449*	0.0166	-0.0390***	-0.1398***	-0.0007	0.0047
	(0.0380)	(0.0254)	(0.0241)	(0.0088)	(0.0312)	(0.0014)	(0.0087)
$\widehat{TotalExpenditure}\times$ Locality Mult. 18%	0.1989***	-0.0094	-0.0050	-0.0499***	-0.1263***	0.0013	0.0023
	(0.0692)	(0.0284)	(0.0395)	(0.0097)	(0.0382)	(0.0010)	(0.0049)
$\widehat{TotalExpenditure}\times$ Locality Mult. 24%	0.1662***	-0.0456***	-0.0378*	-0.0585***	-0.1842***	-0.0002	0.0074**
	(0.0411)	(0.0171)	(0.0196)	(0.0062)	(0.0311)	(0.0008)	(0.0031)
Observations	9,470	9,470	9,470	9,470	9,470	9,470	9,462

Notes: This regression shows results from a regression of real per pupil spending of different categories on an instrument for district expenditure (Post × estimated apportionment under new formula in 2017-2018 minus actual apportionment under old formula in 2017-2018), district fixed effects and year fixed effects. To allow for heterogeneous treatment effects, the estimated total expenditure is interacted with an indicator for if a district had an above median change in estimated per pupil apportionment (Panel A) or the locality multiplier used to determine apportionment under the new funding formula. Standard errors are clustered at the district level. Data from Washington State OSPI. Stars indicate * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Further analysis investigates whether districts with higher locality multipliers disproportionately increased salary spending. Panel B of Table 12 and Table 13 confirms that high-multiplier districts were more likely to increase certificated salary expenditures. The locality multipliers are set by the state, and are meant to serve as a cost of living adjustment for more expensive areas. Since all teacher's unions in the state were able to renegotiate wages before the 2018-2019 school year, those who live in high locality multiplier areas may have been able to use their knowledge of the new state funding formula to negotiate higher wages than their peers. Sun et al. (2024) found that this SFR led school district salary schedules to increase most for experienced teachers which is consistent with Cowen (2009).

Teacher Retention

To assess medium-term retention effects of the SFR, I estimate Eq. (4.2) using retention as the outcome, where $Retention_{d,t}$ is defined as the share of teachers in district d who remained in teaching roles in that district from year $t-1$ to t. Results in Table 14 show small, statistically insignificant effects: a district at the median exposure level experiences a 1 percentage point increase in within-district teacher retention.

Subgroup analysis yields similarly noisy results. When splitting the sample by teacher experience (less than 4 years, more than 15 years, and more than 30 years), none of the coefficients are statistically significant. While the evidence suggests minimal medium-term effects on retention, it does not contradict the short-term findings in Sun et al. (2024). Rather, it suggests that the retention benefits of the SFR may not persist five years after implementation.

TABLE 14. Teacher Retention

	District Retention Rate			
Post $\times$ Δ Est. App.	1.09×10^{-5}**	-2.12×10^{-6}	1.38×10^{-5}	1.62×10^{-5}
	(4.29×10^{-6})	(1.19×10^{-5})	(1.02×10^{-5})	(1.73×10^{-5})
Observations	3,293	3,101	3,193	2,840
Year FE				
District FE				
Sample Cut	All	Exp < 4	Exp > 15	Exp > 30
Observations	3,308	3,116	3,205	2,851

Notes: This table contains the results of a regression of district teacher retention rates on the difference between the estimated apportionment in the 2017-2018 school year under the new funding formula and the actual apportionment in the 2017-2018 school year, a post dummy, their interaction, school year fixed effects, and school district fixed effects. The coefficients are multiplied by 1,000 due to them being small. Standard errors are at the district level. Stars indicate * p <0.1, ** p <0.05, *** p <0.01.

4.6 Conclusion

Given that education represents one of the largest components of public expenditure in the United States, understanding how school districts allocate their funds is critical for both policymakers and researchers. In this paper, I examine how school spending patterns shifted in response to a statewide school finance reform (SFR) in Washington state, prompted by the McCleary Decision. I find that districts responded to changes in funding—particularly those tied to cost-of-living adjustments—by increasing spending on teacher salaries. I also document modest medium-term effects of the reform on teacher retention.

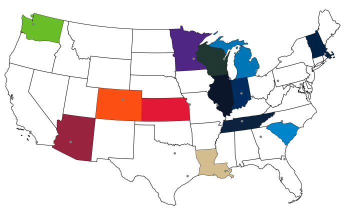
These findings contribute to a growing body of literature providing evidence that increases in school funding can affect how districts prioritize expenditures and support their workforce. Prior research has demonstrated that targeted increases in education spending can improve student outcomes, especially when funding is used for instruction-related expenses such as teacher compensation or reducing class size (Jackson et al., 2016; Lafortune et al., 2018). This study complements that work by highlighting how district-level discretion interacts with reform-induced funding changes.

This analysis also provides insight into the distribution of marginal school expenditures and how exposure to a reform can differentially affect expenditure categories. Understanding these dynamics is particularly important in the context of teacher labor markets. Teacher recruitment and retention remain central concerns for education policymakers, especially in the aftermath of the COVID-19 pandemic, which intensified staffing challenges across many public sectors. In response, school districts have increasingly relied on financial incentives—such as across-the-board raises, signing bonuses, and retention stipends—to retain both early-career and experienced teachers.

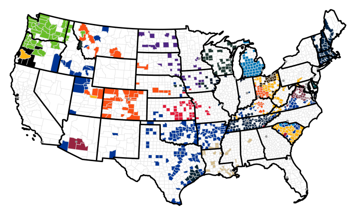
Overall, this research underscores the importance of monitoring not just whether education funds increase, but how those funds are ultimately spent. For reforms to be effective, it is essential that school districts allocate resources in ways that align with policy goals—particularly those aimed at improving teacher quality and workforce stability.

APPENDIX A

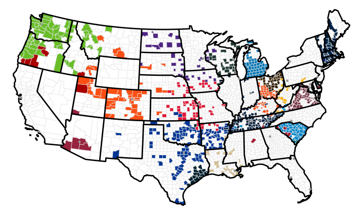
SPORTS BETTING LEGALIZATION AMPLIFIES EMOTIONAL CUES & INTIMATE PARTNER VIOLENCE



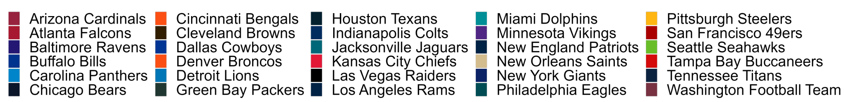
(a) State



(b) Facebook



(c) Seatgeek

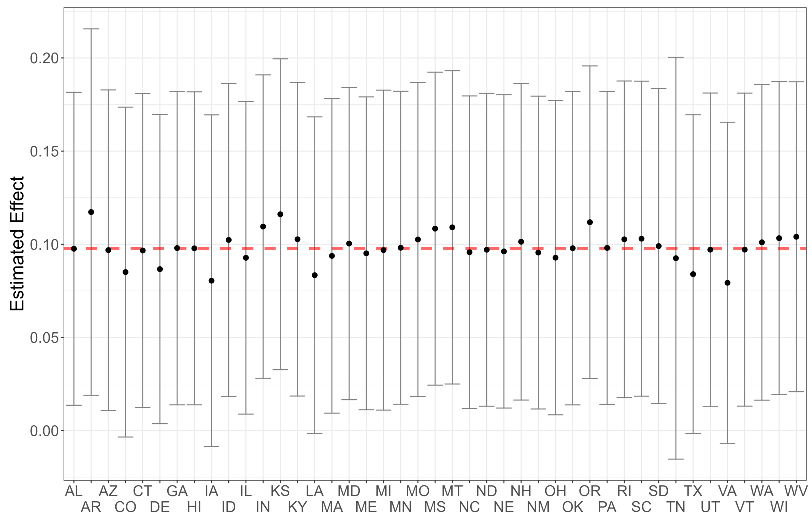


(d) Legend

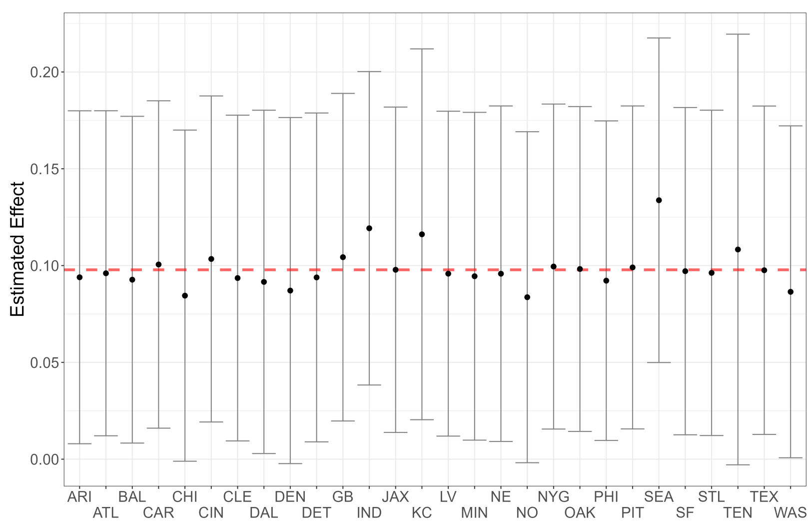
FIGURE A.1. NIBRS Coverage by Home Team

FIGURE A.2. Sensitivity of Betting*Upset Loss Coefficient to Excluding Each State/Team

(a) Exclude Each State



(b) Exclude Each Team



Notes: These figures show sensitivity of our `Betting*Upset Loss` coefficient when we exclude each state or team. The gray bar represents 95% confidence intervals using standard errors clustered at the state-by-team-by-season level. The red dashed line indicates the full-sample treatment effect estimate.

TABLE A.1. Game Outcomes by Teams

Team	**Upset Loss**		**Close Loss**		**Upset Win**	
	Count	% Exp. Win	Count	% Exp. Close	Count	% Exp. Lose
Arizona Cardinals	11	30.6	32	40	13	25
Atlanta Falcons	9	23.1	51	56	9	25.7
Baltimore Ravens	19	28.4	33	44	2	11.8
Buffalo Bills	12	28.6	33	43.4	14	31.1
Carolina Panthers	7	23.3	45	50.6	13	26.5
Chicago Bears	8	25	35	53.8	10	17.2
Cincinnati Bengals	5	14.3	38	42.7	8	18.2
Cleveland Browns	9	36	41	59.4	7	9.5
Dallas Cowboys	10	20	38	49.4	9	33.3
Denver Broncos	7	14	39	54.2	11	28.9
Detroit Lions	12	42.9	43	48.9	10	21.7
Green Bay Packers	14	17.3	30	47.6	5	33.3
Houston Texans	14	31.1	30	45.5	6	11.3
Indianapolis Colts	10	24.4	40	54.1	20	41.7
Jacksonville Jaguars	8	50	47	65.3	14	15.7
Kansas City Chiefs	16	21.1	24	40	8	30.8
Miami Dolphins	7	28	39	45.9	15	26.8
Minnesota Vikings	8	24.2	36	39.1	7	18.9
New England Patriots	16	16.3	24	46.2	3	42.9
New Orleans Saints	22	31.4	27	40.3	9	47.4
New York Giants	6	27.3	36	52.2	22	32.8
New York Jets	1	7.7	41	54.7	17	23.3
OAK/LV Raiders	6	37.5	45	56.2	18	26.9
Philadelphia Eagles	16	27.1	37	52.1	7	26.9
Pittsburgh Steelers	13	23.2	34	42.5	10	43.5
SD/LA Chargers	12	23.5	46	57.5	8	27.6
STL/LA Rams	10	23.8	37	52.9	15	30
San Francisco 49ers	16	27.1	24	46.2	9	19.1
Seattle Seahawks	18	25.7	34	47.2	8	47.1
Tampa Bay Buccaneers	10	33.3	53	58.9	12	25.5
Tennessee Titans	9	28.1	48	54.5	16	34
Washington Commanders	5	31.2	40	48.2	11	19.3

Notes: Data is from the 2011 to 2022 season. Upset loss is defined as losing when the team is expected to win (pre-game spread under -4). Close loss is defined as losing when the team is expected to experience a close matchup (pre-game spread between -4 to 4). Upset win is defined as winning when the team is expected to lose (pre-game spread over 4).

TABLE A.2. Robustness Check of Our Estimates to Various Definitions of Home Team & Geography

	(1)	(2)	(3)	(4)	(5)
Upset Loss	0.0143	0.0206	0.0112	0.0116	0.0155
	(0.0257)	(0.0184)	(0.0182)	(0.0179)	(0.0212)
Upset Loss*Betting	0.1213*	0.1055**	0.0794	0.1145**	0.1321**
	(0.0632)	(0.0493)	(0.0535)	(0.0493)	(0.0616)
Close Loss	-0.0123	-0.0171	-0.0043	-0.0113	-0.0116
	(0.0173)	(0.0126)	(0.0132)	(0.0127)	(0.0160)
Close Loss*Betting	0.0358	0.0538	0.0697	0.0481	0.0578
	(0.0553)	(0.0456)	(0.0456)	(0.0455)	(0.0501)
Upset Win	0.0264	0.0254	0.0146	0.0182	0.0146
	(0.0274)	(0.0203)	(0.0209)	(0.0200)	(0.0240)
Upset Win*Betting	-0.0508	0.0262	0.0086	0.0332	0.0604
	(0.0826)	(0.0682)	(0.0674)	(0.0681)	(0.0785)
Betting	-0.0366	-0.0525	-0.0682**	-0.0591	-0.0260
	(0.0508)	(0.0381)	(0.0348)	(0.0381)	(0.0391)
N	160,653	301,854	299,987	301,854	217,601
Home Team Definition	State	County	County	County	County
Source	N/A	Seatgeek	Facebook	Both	Both
Restrict to Same Fans	No	No	No	No	Yes

Notes: Standard errors are clustered at the season-by-team level. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

TABLE A.3. Robustness Check of Our Estimates

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Upset Loss	0.0303	0.0282	0.0293	0.0246	0.0240	0.0317	0.0140*
	(0.0198)	(0.0201)	(0.0216)	(0.0202)	(0.0273)	(0.0199)	(0.0084)
Upset Loss*Betting	0.1000*	0.1019*	0.1483***	0.1396**	0.1093**	0.1018**	0.0374*
	(0.0511)	(0.0533)	(0.0571)	(0.0580)	(0.0548)	(0.0513)	(0.0225)
Close Loss	-0.0135	-0.0164	-0.0068	-0.0161	-0.0050	-0.0140	-0.0060
	(0.0125)	(0.0126)	(0.0137)	(0.0131)	(0.0162)	(0.0125)	(0.0058)
Close Loss*Betting	0.0764*	0.0700*	0.0857**	0.0771*	0.0677*	0.0808**	0.0325**
	(0.0394)	(0.0388)	(0.0413)	(0.0453)	(0.0395)	(0.0405)	(0.0164)
Upset Win	0.0264	0.0275	0.0183	0.0262	0.0224	0.0283	0.0116
	(0.0199)	(0.0217)	(0.0246)	(0.0204)	(0.0283)	(0.0197)	(0.0096)
Upset Win*Betting	0.0095	0.0181	0.0661	0.0007	0.0185	0.0063	-0.0022
	(0.0668)	(0.0702)	(0.0784)	(0.0841)	(0.0685)	(0.0670)	(0.0298)
Betting	-0.0390	-0.0421	-0.0368	-0.0583	-0.0463	-0.0243	-0.0133
	(0.0349)	(0.0355)	(0.0395)	(0.0424)	(0.0386)	(0.0345)	(0.0150)
N	301,854	284,582	228,545	272,515	187,176	301,854	311,837
Agency Cut?	No	4 Years	8 Years	No	No	No	No
Exclude 2020?	No	No	No	Yes	No	No	No
Sample States?	All	All	All	All	Policy	All	All
Model	Poisson	Poisson	Poisson	Poisson	Poisson	Neg Binom	OLS

Notes: Estimates for columns 1 to 5 are estimated via Poisson model using population as an exposure variable. Estimate for column 6 is estimated via Negative Binomial model using population as an exposure variable. Estimate for column 7 is estimated via OLS using population as the weight variable. Every regression includes controls for agency fixed effects, season fixed effects, week of the season fixed effects, indicator for holiday, average temperature and total precipitation. Standard errors are clustered at the state-by-season-by-team level.Standard errors are clustered at the season-by-team level. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

APPENDIX B

IMPROVING PROGRAMMING IN JUVENILE DETENTION: THE IMPACT OF PROJECT SAFE NEIGHBORHOODS YOUTH OUTREACH FORUMS

TABLE A.4. Intent-to-Treat Effect Estimates, Pooling Both Years

	ITT	Std. Err	95 Percent Conf Int Lower	Upper	Control Mean	Number of Observations	Number of Pods
A. Juvenile Detention Outcomes							
Not Released	-0.02	0.02	-0.05	0.01	0.05	626	86
Days to Release	-3.73	5.73	-14.96	7.49	56.09	600	86
Any New Spells	0.01	0.04	-0.07	0.09	0.47	626	86
Number New Spells	-0.04	0.09	-0.21	0.14	0.84	626	86
Standardized Index	-0.04	0.05	-0.14	0.07	0.01	626	86
B. Arrest Outcomes							
Total	-21.48	19.32	-59.35	16.39	188.16	626	86
Violent	-4.57	5.01	-14.38	5.25	28.74	626	86
Property	-3.04	4.09	-11.07	4.98	16.91	626	86
Drugs	-5.96	5.82	-17.37	5.45	21.50	626	86
Other	-7.91	13.60	-34.56	18.74	121.01	626	86
Standardized Index	-0.07	0.05	-0.17	0.03	0.00	626	86
C. Social Costs of Crime							
Direct Cost, All Crime	-14908	13915	-42181	12365	86737	626	86
Direct Cost, Arrests	-3522	2476	-8375.00	1332	18803	626	86
Willingness to Pay, All Crime	-181257**	66572	-311738	-50776	420642	626	86
Willingness to Pay, Arrests	-23230**	8551	-39991	-6470	56362	626	86
D. Education Outcomes							
Attended any Days	0.01	0.03	-0.06	0.08	0.56	611	84
Number of Days	4.88	2.42	0.13	9.62	24.76	611	84
Has GPA	0.076*	0.03	0.01	0.14	0.47	611	84
GPA	-0.01	0.12	-0.25	0.23	1.31	295	84
Standardized Index	0.090*	0.04	0.00	0.18	-0.06	611	85

Notes: This table shows estimates of the intent-to-treat effect of the program, pooling both the first and second years. Random assignment is only correlated with participation in the second year of the study. Standard errors clustered on individual and residential pod using the method described in Cameron and Miller (2015). The pooled sample includes 86 pod-lottery clusters. All regressions conditional on block fixed effects and the baseline covariates described in section . Stars indicate: * $p<0.1$, ** $p<0.05$.

TABLE A.5. Intent-to-Treat Effect Estimates, First Year Only

	ITT	Std. Err	95 Percent Conf Int Lower	Upper	Control Mean	Number of Observations	Number of Pods
A. Juvenile Detention Outcomes							
Not Released	-0.02	0.01	-0.04	0.00	0.02	373	56
Days to Release	-6.40	7.65	-21.39	8.59	70.47	369	56
Any New Spells	0.06	0.05	-0.04	0.17	0.39	373	56
Number New Spells	0.10	0.11	-0.12	0.32	0.67	373	56
Standardized Index	0.02	0.05	-0.09	0.12	0.01	373	56
B. Arrest Outcomes							
Total	-10.21	25.73	-60.64	40.22	174.90	373	56
Violent	3.17	5.92	-8.43	14.76	24.69	373	56
Property	-0.59	5.24	-10.86	9.67	14.23	373	56
Drugs	-4.86	9.53	-23.53	13.81	23.85	373	56
Other	-7.92	18.03	-43.26	27.41	112.13	373	56
Standardized Index	-0.03	0.07	-0.17	0.12	0.00	373	56
C. Social Costs of Crime							
Direct Cost, All Crime	4335	14235	-23565	32235	69492	373	56
Direct Cost, Arrests	1052	2530	-3908	6011	15085	373	56
Willingness to Pay, All Crime	-33531	70814	-172327	105265	312354	373	56
Willingness to Pay, Arrests	-3333	8765	-20513	13847	41806	373	56
D. Education Outcomes							
Attended any Days	-0.04	0.04	-0.12	0.04	0.60	367	56
Number of Days	4.24	2.90	-1.45	9.93	27.01	367	56
Has GPA	0.03	0.04	-0.06	0.12	0.52	367	56
GPA	-0.21	0.14	-0.48	0.05	1.43	190	56
Standardized Index	0.00	0.05	-0.10	0.10	-0.06	367	55

Notes: This table shows estimates of the intent-to-treat effect for the first year of the program when random assignment is uncorrelated with program participation. These results are shown for completeness. The first year includes 56 pod-lottery clusters. Standard errors clustered on individual and residential pod using the method described in Cameron and Miller (2015). All regressions are based on Equation 3.2 using all 626 youth in the study. The year two observations improve the precision of the estimated coefficients on baseline covariates but do not otherwise contribute to the treatment effect estimates. Stars indicate: * $p<0.1$, ** $p<0.05$.

APPENDIX C

DO BUDGETS REFLECT PRIORITIES? THE IMPACT OF SCHOOL FUNDING REFORM ON DISTRICTS AND TEACHERS IN WASHINGTON STATE

C.1 Calculation of Apportionment

Apportionment for districts can be calculated in the following way:

1. Calculate staff units from the following equations:

To calculate certificated instructional staff allocation:

$$\begin{aligned}
TeacherUnits_d &= \sum_{g=k}^{12} \frac{Enrollment_{g,d}}{ClassSize_g} * (1 + PlanningTimeFactor_g) \\
TeacherUnits_d &= \sum_{g=k}^{3} \frac{Enrollment_{g,d}}{17} * (1 + 0.155) + \sum_{g=4}^{6} \frac{Enrollment_{g,d}}{27} * (1 + 0.155) \\
&+ \sum_{g=7}^{8} \frac{Enrollment_{g,d}}{28.53} * (1 + 0.2) + \sum_{g=9}^{12} \frac{Enrollment_{g,d}}{28.73} * (1 + 0.2) \\
&+ \frac{Enrollment_{middleschoolCTE,d}}{23} * (1 + 0.2) \\
&+ \frac{Enrollment_{highschoolCTE,d}}{23} * (1 + 0.2) \\
&+ \frac{Enrollment_{skillscenter}}{20} * (1 + 0.2) + \frac{Enrollment_{labscience,d}}{19.98} * (1 + 0.2)
\end{aligned} \tag{C.1}$$

To calculate certificated administrative staff (principals):

$$\begin{aligned}
PrincipalUnits_d &= 1.253 * \frac{ElementryEnrollment_d}{400} + 1.353 * \frac{MiddleEnrollment_d}{432} \\
&+ 1.88 * \frac{HighEnrollment_d}{600}
\end{aligned} \tag{C.2}$$

To calculate other certificated instructional staff units:

$$CISUnits_d = 1.291 * \frac{ElementryEnrollment_d}{400} + 1.803 * \frac{MiddleEnrollment_d}{432} + 13.189 * \frac{HighEnrollment_d}{600} \tag{C.3}$$

To calculate classified staff units:

$$CLSUnits_d = 4.7665 * \frac{ElementryEnrollment_d}{400} + 5.059 * \frac{MiddleEnrollment_d}{432} + 7.027 * \frac{HighEnrollment_d}{600} \tag{C.4}$$

To calculate units that are shared amongst students in the district (technology, facilities, maintenance, mechanics, laborers):

$$DistrictWideSupportUnits_d = 2.773 * \frac{TotalEnrollment_d}{1000} \tag{C.5}$$

To calculate central district administration units, take 5.3% of the district's total employees of the district allocation, Of this, 25.47% will be certificated administrative staff and 74.53% will be classified administrative staff:

$$
\begin{aligned}
CentralAdministrationCertificated_d = 0.0134991 * (&TeacherUnits_d \\
&+ PrincipalUnits_d \\
&+ CISUnits_d + CLSUnits_d \\
&+ DistrictWideSupportUnits_d) \\
CentralAdministrationClassified_d = 0.0395009 * (&TeacherUnits_d \\
&+ PrincipalUnits_d \\
&+ CISUnits_d + CLSUnits_d \\
&+ DistrictWideSupportUnits_d)
\end{aligned}
\tag{C.6}
$$

2. Calculate Materials, Supplies, and Operating Costs

To calculate Materials, Supplies, and Operating Costs (MSOC):

$$
\begin{aligned}
MSOC_d = 1264.07 * (ElementryEnrollment_d + MiddleEnrollment_d) \\
+ 1437.71 * HighEnrollment_d + 1495.56 * CTEandSkillCenterEnrollment_d
\end{aligned}
\tag{C.7}
$$

3. Calculate Payroll Taxes and Benefits Allocation:

Payroll taxes and benefits are funded by the state at a rate of 23.70% for certificated salary maintenance, 23.06% for certificated salary increase, 24.70% for classified salary maintenance and 21.20% for classified salary increase.

4. Calculate Insurance Benefits:

School districts receive $10,127.64 per certificated unit for certificated staff and 1.152*10,127.64 ($11,667.04) for classified staff insurance benefits.

$$
\begin{aligned}
Insurance_d =& 10127.64 * (TeacherUnits_d + PrincipalUnits_d + CISUnits_d \\
& + CentralAdminstrationCertificated_d) \\
& + 11667.04 * (CLSUnits_d + DistrictWideSupportUnits_d \\
& + CentralAdministrationClassified_d)
\end{aligned} \tag{C.8}
$$

5. Calculate Substitute Teacher Pay:

Every full teacher FTE is allocated four substitute days at a rate of \$151.86 per day. Therefore, the calculation is:

$$
Substitute_d = 4 * 151.86 * TeacherUnits_d \tag{C.9}
$$

6. Calculate Local Deductible Revenues

Certain local revenues are deductible from the apportionment calculation. These include local in lieu of taxes including county in lieu of tax payments by housing authorities or from lands purchased by the Department of Natural Resources, and Federal in lieu of taxes including revenue from the Federal Housing Administration, Bureau of Land Management, military forest yield pursuant to Public Law 9799, and reclamation projects.

$$
LocalDeductibleRevenues_d = LocalinLieuofTaxes_d + FederalinLieuofTaxes_d \tag{C.10}
$$

7. Calculate Fire District Payments

Fire district payments reimburse eligible districts at a rate of \$1.145926 per pupil.

$$FireDistrict_d = Eligibility_d * 1.145926 * TotalEnrollment_d \tag{C.11}$$

8. Calculate Apportionment

The apportionment calculation is the sum of the employee allocations times their respective state salary allocation amounts and a district level multiplier, the MSOC, an adjustment for small school and small districts, payroll taxes and benefits, insurance benefits, substitute teacher pay, fire district revenues, and minus local deductible revenues:

$$\begin{aligned} Apportionment_d = {} & locality_d * (65,215 * TeacherUnits_d + 96,805 * PrincipalUnits_d \\ & + 65,215 * CISUnits_d + 46,784.33 * CLSUnits_d \\ & + 46,784.33 * DistrictWideSupportUnits_d \\ & + 96,805 * CentralAdministratioCertificated_d \\ & + 46,784.33 * CentralAdministrationClassified_d) + MSOC_d \\ & + SmallSchoolandDistrictAdj_d + PayrollTax_d + Insurance_d \\ & + Substitute_d + FireDistrict_d - LocalDeductibleRevenues_d \end{aligned} \tag{C.12}$$

TABLE A.6. Marginal Change in Per Pupil Spending Using Exposure and Fixed Effects

Panel A	Cert. Salary	Clas. Salary	Benefits	Supplies	Services	Travel	Cap. Outlay
Post × Δ Expenditure 2017/18 to 2018/19	0.0566	0.0334	-0.0229	-0.0036	0.0432	0.0013	0.0055
	(0.0656)	(0.0262)	(0.0287)	(0.0169)	(0.0545)	(0.0011)	(0.0063)
Observations	9,701	9,701	9,701	9,701	9,701	9,704	9,697
Panel B							
Post × Δ Actual Apportionment	0.3759*	0.1124	0.2699**	-0.0221	0.2799**	-7.06×10^{-5}	-0.0358**
	(0.2187)	(0.0885)	(0.1296)	(0.0339)	(0.1188)	(0.0040)	(0.0164)
Observations	9,727	9,727	9,727	9,727	9,727	9,727	9,719
Panel C							
Post × Δ Est. Apportionment	0.4378***	0.1521**	0.2868***	-0.0041	0.2382***	-0.0012	-0.0340***
	(0.1680)	(0.0643)	(0.1003)	(0.0244)	(0.0899)	(0.0029)	(0.0127)
Observations	9,701	9,701	9,701	9,701	9,701	9,701	9,693
Year FE							
District FE							

Notes: This table shows results from a regression of real per pupil spending of different categories on different measures of exposure to the policy change interacted with a post-dummy, year fixed effects and school district fixed effects. Panel A defines exposure to the policy change as the change in expenditure from the 2017-2018 to 2018-2019 school year. Panel B defines exposure to the policy change as the change in apportionment by school district between the 2017-2018 and 2018-2019 school years. Panel C defines exposure to the policy change as the difference between apportionment under the new funding formula for 2017-2018 district characteristics and the actual apportionment for 2017-2018. Standard errors are clustered at the district level. Stars indicate * p <0.1, ** p <0.05, *** p <0.01.

TABLE A.7. Marginal Per Pupil Spending Using Exposure and IV

Panel A	Total Expenditure	Cert. Salary	Clas. Salary	Benefits	Supplies	Services	Travel	Cap. Outlay
$\widehat{TotalExpenditure}$	0.5017	0.2955	-0.2025	-0.0319	0.3824	0.0107	0.0443	
	(0.3774)	(0.2699)	(0.4897)	(0.1784)	(0.4610)	(0.0140)	(0.0776)	
Observations	9,701	9,701	9,701	9,701	9,701	9,701	9,693	
Panel B								
$\widehat{TotalExpenditure}$	0.3835***	0.1146**	0.2753***	-0.0225	0.2855**	-7.2×10^{-5}	-0.0359	
	(0.1050)	(0.0551)	(0.0618)	(0.0391)	(0.1445)	(0.0040)	(0.0231)	
Observations	9,727	9,727	9,727	9,727	9,727	9,727	9,719	
Panel C								
$\widehat{TotalExpenditurePerPupil}$	0.4070***	0.1414***	0.2666***	-0.0038	0.2214**	-0.0011	-0.0316**	
	(0.0768)	(0.0394)	(0.0419)	(0.0229)	(0.0946)	(0.0027)	(0.0151)	
Observations	9,701	9,701	9,701	9,701	9,701	9,701	9,693	
R^2	0.92606	0.91896	0.92340	0.62496	0.87529	0.50692	0.01495	
Within R^2	0.78473	0.71316	0.72944	-0.04911	0.62890	-0.04669	-0.28668	
Year FE								
District FE								

Notes: This table shows results from a regression of real per pupil spending of different categories on an instrument for district spending (Post × a measure of exposure to the policy change approximately differently across panels), district fixed effects and year fixed effects. Panel A defines exposure to the policy change as the change in expenditure from the 2017-2018 to 2018-2019 school year. Panel B defines exposure to the policy change as the change in apportionment by school district between the 2017-2018 and 2018-2019 school years. Panel C defines exposure to the policy change as the difference between apportionment under the new funding formula for 2017-2018 district characteristics and the actual apportionment for 2017-2018. Standard errors are clustered at the district level. Data from Washington State OSPI. Stars indicate * p <0.1, ** p <0.05, *** p <0.01.

REFERENCES CITED

Adams-Prassl, A., Huttunen, K., Nix, E., and Zhang, N. (2023). The dynamics of abusive relationships.

Aizer, A. (2010). The gender wage gap and domestic violence. *American Economic Review*, 100(4):1847–1859.

Aizer, A. and Doyle, J. J. (2015). Juvenile Incarceration, Human Capital, and Future Crime: Evidence from Randomly Assigned Judges. *Quarterly Journal of Economics*, 130(2):1–46.

American Gaming Association (2023). Assessing shifts in the sports betting market 5 years post-paspa. `https://www.americangaming.org/wp-content/uploads/2023/05/AGA_PASPA_LSBResearch.pdf`.

Anderberg, D., Rainer, H., Wadsworth, J., and Wilson, T. (2016). Unemployment and domestic violence: Theory and evidence. *The Economic Journal*, 126(597):1947–1979.

Angrist, J. D., Imbens, G. W., and Rubin, D. B. (1996). Identification of causal effects using instrumental variables. *Journal of the American statistical Association*, 91(434):444–455.

Arenas-Arroyo, E., Fernandez-Kranz, D., and Nollenberger, N. (2021). Intimate partner violence under forced cohabitation and economic stress: Evidence from the covid-19 pandemic. *Journal of Public Economics*, 194:104350–104350.

Baker, S., Balthrop, J., Johnson, M. J., Kotter, J. D., and Pisciotta, K. (2024). Gambling away stability: Sports betting's impact on vulnerable households. *SSRN Working Paper No. 4881086*.

Beck, A., Cantor, D., Hartge, J., and Smith, T. (2013). Sexual Victimization in Juvenile Facilities Reported by Youth, 2012. Technical Report June, U.S. Department of Justice, Office of Justice Programs, Bureau of Justice Statistics.

Becker, G. S. (1968). Crime and punishment: An economic approach. *Journal of Political Economy*, 76(2):169–217.

Beland, L.-P. and Brent, D. A. (2018). Traffic and crime. *Journal of Public Economics*, 160:96–116.

Benjamin, D. J., Berger, J. O., Johannesson, M., Nosek, B. A., Wagenmakers, E.-J., Berk, R., Bollen, K. A., Brembs, B., Brown, L., Camerer, C., et al. (2018). Redefine statistical significance. *Nature Human Behaviour*, 2(1):6–10.

Bertrand, M., Duflo, E., and Mullainathan, S. (2004). How much should we trust differences-in-differences estimates? *The Quarterly Journal of Economics*, 119(1):249–275.

Bhuller, M., Dahl, G. B., Løken, K. V., and Mogstad, M. (2022). Domestic violence and the mental health and well-being of victims and their children. Technical report, National Bureau of Economic Research.

Booher-Jennings, J. (2005). Below the bubble: "educational triage" and the texas accountability system. *American educational research journal*, 42(2):231–268.

Bottoms, A. and Tankebe, J. (2012). Beyond procedural justice: A dialogic approach to legitimacy in criminal justice. *Journal of Criminal Law and Criminology*, 102:119–170.

Braga, A. A. and Kennedy, D. M. (2021). *A framework for addressing violence and serious crime: Focused deterrence, legitimacy, and prevention*. Cambridge University Press.

Braga, A. A., Kennedy, D. M., Waring, E. J., and Piehl, A. M. P. (2001). Problem-oriented policing, deterrence, and youth violence: An evaluation of boston's operation ceasefire. *Journal of Research in Crime Delinquency*, 38:195–226.

Braga, A. A., Weisburd, D., and Turchan, B. (2018). Focused deterrence strategies and crime control: An updated systematic review and meta-analysis of the empirical evidence. *Criminology & Public Policy*, 17(1):205–250.

Brooks, C. D. and Springer, M. G. (2024). Esser-ting preferences: Examining school district preferences for using federal pandemic relief fundings. *EdWorkingPaper: 24-913*.

Cameron, A. C. and Miller, D. L. (2015). A practitioner's guide to cluster-robust inference. *Journal of Human Resources*, 50(2):317–372.

Can, E., Nichols, M. W., and Pavlopoulos, V. (2023). The effects of sports betting on casino gambling and lottery. *SSRN Working Paper No. 4659440*.

Candelaria, C. A. and Shores, K. A. (2019). Court-ordered finance reforms in the adequacy era: Heterogeneous causal effects and sensitivity. *Education finance and policy*, 14(1):31–60.

Card, D. and Dahl, G. B. (2011). Family violence and football: The effect of unexpected emotional cues on violent behavior. *The Quarterly Journal of Economics*, 126(1):103–143.

Card, D. and Payne, A. (2002). School finance reform, the distribution of school spending, and the distribution of student test scores. *Journal of public economics*, 83(1):49–82.

Cardazzi, A., McCannon, B. C., Humphreys, B. R., and Rodriguez, Z. (2022). Emotional cues and violent behavior: Unexpected basketball losses increase incidents of family violence. *Journal of Law, Economics, & Organization*.

Carr, J. B. and Packham, A. (2021). Snap schedules and domestic violence. *Journal of Policy Analysis and Management*, 40(2):412–452.

Carrell, S. E. and Hoekstra, M. (2012). Family business or social problem? the cost of unreported domestic violence. *Journal of Policy Analysis and Management*, 31(4):861–875.

Carrell, S. E., Hoekstra, M., and Kuka, E. (2018). The long-run effects of disruptive peers. *American Economic Review*, 108(11):3377–3415.

Carrell, S. E. and Hoekstra, M. L. (2010). Externalities in the classroom: How children exposed to domestic violence affect everyone's kids. *American Economic Journal: Applied Economics*, 2(1):211–228.

Centers for Disease Control and Prevention (2024). About intimate partner violence. `https://www.cdc.gov/intimate-partner-violence/about/index.html`.

Chaisemartin, C. D. and d'Haultfoeuille, X. (2020). Two-way fixed effects estimators with heterogeneous treatment effects. *American Economic Review*, 110:2964–2996.

Clotfelter, C. T., Ladd, H. F., and Vigdor, J. L. (2006). Teacher-student matching and the assessment of teacher effectiveness. *The Journal of human resources*, XLI(4):778–820.

Cohen, M. A. and Piquero, A. R. (2009). New Evidence on the Monetary Value of Saving a High Risk Youth. *Journal of Quantitative Criminology*, 25:25–49.

Cohen, M. A., Rust, R. T., Steen, S., and Tidd, S. T. (2004). Willingness to pay for crime control programs. *Criminology*, 42(1):89–110.

Collins, C. (2022). Emotional cues, election outcomes, and intimate partner violence. *SSRN Working Paper No. 4278930*.

Corsaro, N., Hunt, E. D., Hipple, N. K., and McGarrell, E. F. (2012). The impact of drug market pulling levers policing on neighborhood violence: An evaluation of the high point drug market intervention. *Criminology and Public Policy*, 11(2):167–199.

Cowen, J. M. (2009). Teacher unions and teacher compensation: New evidence for the impact of bargaining. *Journal of education finance*, 35(2):172–193.

Davis, J. M. and Heller, S. B. (2020). Rethinking the benefits of youth employment programs: The heterogeneous effects of summer jobs. *Review of Economics and Statistics*, 102(4):664–677.

Dee, T. S., Jacob, B., and Schwartz, N. L. (2013). The effects of nclb on school resources and practices. *Educational evaluation and policy analysis*, 35(2):252–279.

Dobkin, C. and Puller, S. L. (2007). The effects of government transfers on monthly cycles in drug abuse, hospitalization and mortality. *Journal of Public Economics*, 91(11-12):2137–2157.

Dowling, N., Suomi, A., Jackson, A., Lavis, T., Patford, J., Cockman, S., Thomas, S., Bellringer, M., Koziol-Mclain, J., Battersby, M., Harvey, P., and Abbott, M. (2016). Problem gambling and intimate partner violence: A systematic review and meta-analysis. *Trauma, Violence & Abuse*, 17(1):43–61.

Edelstein, J. (2021). One extra week of nfl games, millions of extra dollars for sportsbooks. Sports Handle. `https://sportshandle.com/extra-nfl-week-sportsbooks/`.

Farrington, D. P. and Welsh, B. C. (2005). Randomized experiments in criminology: What have we learned in the last two decades? *Journal of Experimental Criminology*, 1:9–38.

Gandar, J., Zuber, R., O'brien, T., and Russo, B. (1988). Testing rationality in the point spread betting market. *The Journal of Finance*, 43(4):995–1008.

Goodman-Bacon, A. (2021). Difference-in-differences with variation in treatment timing. *Journal of Econometrics*.

Grissom, J. A., Kalogrides, D., and Loeb, S. (2017). Strategic staffing? how performance pressures affect the distribution of teachers within schools and resulting student achievement. *American educational research journal*, 54(6):1079–1116.

Grommon, E., Carson, D. C., and Kenney, L. (2018). An experimental trial of a dog-training program in a juvenile detention center. *Journal of Experimental Criminology*.

Heller, S. B., Shah, A. K., Guryan, J., Ludwig, J., Mullainathan, S., and Pollack, H. A. (2017). Thinking, Fast and Slow? Some Field Experiments to Reduce Crime and Dropout in Chicago. *Quarterly Journal of Economics*, 132(1):1–54.

Hendricks, M. D. (2014). Does it pay to pay teachers more? evidence from texas. *Journal of public economics*, 109:50–63.

Hing, N., O'Mullan, C., Nuske, E., Breen, H., Mainey, L., Taylor, A., Greer, N., Jenkinson, R., Thomas, A., Lee, J., et al. (2022). Gambling-related intimate partner violence against women: A grounded theory model of individual and relationship determinants. *Journal of Interpersonal Violence*, 37(19-20):NP18639–NP18665.

Hollenbeck, B., Larsen, P., and Proserpio, D. (2024). The financial consequences of legalized sports gambling. *SSRN Working Paper No. 4881086*.

Huble, C. (2023). Ncpg statement on the betting on our future act. `https://www.ncpgambling.org/news/ncpg-statement-on-the-betting-on-our-future-act/`.

Humphreys, B. R. (2021). Legalized sports betting, vlt gambling, and state gambling revenues: Evidence from west virginia. *Eastern Economic Journal*, 47(1):9–28.

Jackson, C. K. (2018). Does school spending matter? the new literature on an old question. *NBER Working Paper Series*, pages 25368–.

Jackson, C. K., Johnson, R. C., and Persico, C. (2016). The effects of school spending on educational and economic outcomes: Evidence from school finance reforms. *The Quarterly journal of economics*, 131(1):157–218.

Jackson, C. K. and Mackevicius, C. (2021). The distribution of school spending impacts. *NBER Working Paper Series*.

Jackson, C. K., Wigger, C., and Xiong, H. (2021). Do school spending cuts matter? evidence from the great recession. *American economic journal. Economic policy*, 13(2):304–335.

Kennedy, D. M. (1997). Pulling levers: Chronic offenders, high-crime settings, and a theory of prevention. *Valparaiso University Law Review*, 31(2):449–484.

Kennedy, D. M. (2006). Old wine in new bottles: Policing and the lessons of pulling levers. In Weisburd, D. L. and Braga, A., editors, *Police Innovation: Contrasting Perspectives*. Cambridge University Press, New York.

Kennedy, D. M. (2009). Drugs, race and common ground: Reflections on the high point intervention. *NIJ Journal*, 262:12–17.

Kimbrell, C. S., Wilson, D. B., and Olaghere, A. (2023). Restorative justice programs and practices in juvenile justice: An updated systematic review and meta-analysis for effectiveness. *Criminology & Public Policy*, 22(1):161–195.

Korman, L. M., Collins, J., Dutton, D., Dhayananthan, B., Littman-Sharp, N., and Skinner, W. (2008a). Problem gambling and intimate partner violence. *Journal of Gambling Studies*, 24:13–23.

Korman, L. M., Collins, J., Dutton, D., Dhayananthan, B., Littman-Sharp, N., and Skinner, W. (2008b). Problem gambling and intimate partner violence. *Journal of Gambling Studies*, 24:13–23.

Kreisman, D. and Steinberg, M. P. (2019). The effect of increased funding on student achievement: Evidence from texas's small district adjustment. *Journal of public economics*, 176:118–141.

Lafortune, J., Rothstein, J., and Schanzenbach, D. W. (2018). School finance reform and the distribution of student achievement. *American economic journal. Applied economics*, 10(2):1–26.

Landenberger, N. A. and Lipsey, M. W. (2005). The positive effects of cognitive–behavioral programs for offenders: A meta-analysis of factors associated with effective treatment. *Journal of Experimental Criminology*, 1(4):451–476.

Lee, D. S. and McCrary, J. (2017). The deterrence effect of prison: Dynamic theory and evidence. *Advances in Econometrics*, 38.

Lin, T.-C. and Pursiainen, V. (2023). The disutility of stock market losses: Evidence from domestic violence. *The Review of Financial Studies*, 36(4):1703–1736.

Lindo, J. M., Schaller, J., and Hansen, B. (2018). Caution! men not at work: Gender-specific labor market conditions and child maltreatment. *Journal of Public Economics*, 163:77–98.

Lindo, J. M., Swensen, I. D., and Waddell, G. R. (2022). Effects of violent media content: Evidence from the rise of the ufc. *Journal of Health Economics*, 83:102623.

McGarrell, E. F., Kroovand Hipple, N., Corsaro, N., Bynum, T. S., Perez, H., Zimmermann, C. A., and Garmo, M. (2009). *Project Safe Neighborhoods - A National Program to Reduce Gun Crime: Final Project Report*. US Department of Justice, Office of Justice Programs, National Institute of Justice.

Miller, T. R., Cohen, M. A., and Wiersema, B. (1996). Victim Costs and Consequences: A New Look. Technical report, U.S. Department of Justice, Office of Justice Programs, National Institute of Justice.

Nagin, D. S. (2013). Deterrence in the twenty-first century. *Crime and Justice*, 42(1):199–263.

National Domestic Violence Hotline (2023). Domestic violence statistics. `https://thehotline.org/stakeholders/domestic-violence-statistics/`.

Nizalova, O. Y. and Murtazashvili, I. (2016). Exogenous treatment and endogenous factors: Vanishing of omitted variable bias on the interaction term. *Journal of Econometric Methods*, 5(1):71–77.

Oxford Economics (2017). Economic impact of legalized sports betting. `https://www.americangaming.org/wp-content/uploads/2018/12/AGA-Oxford-Sports-Betting-Economic-Impact-Report1-1.pdf`.

Pankoff, L. D. (1968). Market efficiency and football betting. *The Journal of Business*, 41(2):203–214.

Papachristos, A. V. and Kirk, D. S. (2015). Changing the street dynamic: Evaluating chicago's group violence reduction strategy. *Criminology & Public Policy*, 14.

Papachristos, A. V., Meares, T. L., and Fagan, J. (2007). Attention Felons: Evaluating Project Safe Neighborhoods in Chicago. *Journal of Empirical Legal Studies*, 4(2):223–272.

Peterson, C., Kearns, M. C., McIntosh, W. L., Estefan, L. F., Nicolaidis, C., McCollister, K. E., Gordon, A., and Florence, C. (2018). Lifetime economic burden of intimate partner violence among u.s. adults. *American Journal of Preventive Medicine*, 55(4):433–444.

Petteruti, A., Schindler, M., and Ziedenberg, J. (2014). Sticker Shock: Calculating the Full Price Tag for Youth Incarceration. Technical report, Justice Policy Institute.

Raphael, S. and Ludwig, J. (2003). Prison sentence enhancements: The case of project exile. *Evaluating gun policy: Effects on crime and violence*, 251:274–277.

Schneider, D., Harknett, K., and McLanahan, S. (2016). Intimate partner violence in the great recession. *Demography*, 53(2):471–505.

Shem-Tov, Y., Raphael, S., and Skog, A. (2021). Can restorative justice conferencing reduce recidivism? evidence from the make-it-right program. *NBER Working Paper Series*.

Sickmund, M., Sladky, T., Kang, W., and Puzzanchera, C. (2019). Easy Access to the Census of Juveniles in Residential Placement. https://www.ojjdp.gov/ojstatbb/ezacjrp/.

Springer, M. G., Liu, K., and Guthrie, J. W. (2009). The impact of school finance litigation on resource distribution: a comparison of court-mandated equity and adequacy reforms. *Education economics*, 17(4):421–444.

Stephens Jr, M. (2006). Paycheque receipt and the timing of consumption. *The Economic Journal*, 116(513):680–701.

Sun, L. and Abraham, S. (2021). Estimating dynamic treatment effects in event studies with heterogeneous treatment effects. *Journal of Econometrics*, 225:175–199.

Sun, M., Candelaria, C. A., Knight, D., LeClair, Z., Kabourek, S. E., and Chang, K. (2024). The effects and local implementation of school finance reforms on teacher salary, hiring, and turnover. *Educational evaluation and policy analysis*.

Suomi, A., Jackson, A. C., Dowling, N. A., Lavis, T., Patford, J., Thomas, S. A., Harvey, P., Abbott, M., Bellringer, M. E., Koziol-McLain, J., et al. (2013). Problem gambling and family violence: family member reports of prevalence, family impacts and family coping. *Asian Journal of Gambling Issues and Public Health*, 3(1):1–15.

Tyler, T. R. (2006a). Psychological perspectives on legitimacy and legitimation. *Annual Review of Psychology*, 57:375–400.

Tyler, T. R. (2006b). *Why People Obey the Law*. Princeton University Press.

Tyler, T. R. and Huo, Y. J. (2002). *Trust in the Law: Encouraging Public Cooperation with the Police and Courts*. Russell Sage Foundation.

Tyler, T. R. and Trinkner, R. (2018). *Why Children Follow Rules: Legal Socialization and the Development of Legitimacy*. Oxford University Press.

UN Women (2023). Facts and figures: Ending violence against women. `https://www.unwomen.org/en/what-we-do/ending-violence-against-women/facts-and-figures`.

Urban Institute (2021). *Project: State and Local Backgrounders*.

van der Maas, M., Cho, S. R., and Nower, L. (2022). Problem gambling message board activity and the legalization of sports betting in the us: A mixed methods approach. *Computers in Human Behavior*, 128:107133–.

Vooren, M., Rud, I., Cornelisz, I., Van Klaveren, C., Groot, W., and van den Brink, H. M. (2022). The effects of a restorative justice programme (halt) on educational outcomes and recidivism of young people. *Journal of Experimental Criminology*.

Walsh, A. (2023). Which is the sport americans bet on most? Rotowire. `https://www.rotowire.com/article/us-sports-betting-news-which-is-the-sport-americans-bet-on-most-71608`.

Wasserstein, R. L. and Lazar, N. A. (2016). The asa statement on p-values: context, process, and purpose.

Western, B. and Pettit, B. (2010). *Collateral Costs: Incarceration's Effect on Economic Mobility*. The Pew Charitable Trusts.

Zamarro, G., Camp, A., McGee, J., Wilson, T., and Vernon, M. (2024). Teacher salary raises and turnover: Evidence from the first year of the arkansas learns act. *EdWorking Paper*, (1):1–32.