Gender, Justice, and Crime: An Empirical Analysis

By

Jinglin Wen

A Thesis Submitted for the Degree of

Doctor of Philosophy in Economics

Adam Smith Business School

College of Social Sciences

University of Glasgow

January 2023

ABSTRACT

This thesis consists of five chapters relating to gender justice, education, sanction law, and public policy.

In the first chapter, I examine the effect of female leadership in local government on violence against women. Regression discontinuity estimates show that elections of female mayors decrease violence specifically targeted toward women but leave other violence unaffected. Moreover, the study explores behavioral responses by victims. Evidence suggests that female victims are more likely to report violence against them after female mayors take office. Importantly, female victories are followed by greater police responsiveness to violence against women.

In the second chapter, I study the impact of the appointment of female chief officers in policing on female salient crimes. Evidence shows that appointing more female chief officers leads to a significant increase in recorded female salient crimes. Yet, this rise is good news, which is due not to a rise in actual crimes committed but, rather, to greater reporting or recording. I also find that the appointment of female chief officers decreases violence against women.

The third chapter, joint with Prateek Chandra Bhan, studies the role modelling effect of historical statues. We conduct a Ramdomised Controlled Trail (RCT) and find that treatment with a virtual tour to four statues of distinct role models in Jaipur leads to an increase in students' academic performance. Evidence from heterogeneity analysis suggests that the statues intervention affects only boys.

In the fourth chapter, I study the effect of health care on domestic violence. I find that improving access to medical care significantly decreases domestic violence against women. I corroborate this by using an instrument variable approach.

I close by examining the effect of expanding Medicaid coverage. Exploration of mechanisms suggests that expanding health insurance may save battered women's lives by increasing economic independence for women.

In the fifth chapter, I look into the impact of the severity of punishment on the behavior of victims, police, and potential criminals. Using multiple causal identification strategies, I study the felony theft thresholds. Evidence shows that the felony theft thresholds have significant effect on crime reporting: felony thefts are more likely to be reported to the police by victims. Moreover, raising the felony thresholds reduces deterrence, and therefore leads to a decline in law and order.

Contents

ABSTRACT **iii**

ACKNOWLEDGEMENTS **xiv**

DECLARATION **xvi**

INTRODUCTION **xvii**

1 Female Mayors and Violence against Women: Evidence from the U.S. **1**

1.1 Introduction 1

1.2 Mechanisms 6

1.3 Data 9

1.4 Effects of Female Mayors on Homicides 12

1.5 Effects on Police Action and Crime Reporting 23

1.6 Discussion 28

1.7 Conclusion 30

2 Female Chief Officers and Crime: Evidence from England and Wales **57**

2.1 Introduction 57

2.2 Background 63

2.3 Data 66

2.4 Analysis of the Appointment of Female Chief Officers 68

2.5 Discussion 78

2.6 Conclusion 79

3 Inspirations and Gender Disparities Set in Stones **95**

3.1 Introduction 96
3.2 Background 100
3.3 Research Design 102
3.4 Data and Estimation Strategy 107
3.5 Results 112
3.6 Discussion 118
3.7 Conclusion 126

4 Does Health Care Save Battered Women's Lives: Evidence from the U.S. 148

4.1 Introduction 149
4.2 Data and Initial Descriptive Analysis 153
4.3 Empirical Strategy 156
4.4 Results 160
4.5 Medicaid Expansion 165
4.6 Mechanisms 172
4.7 Conclusion 176

5 Punishment and Crime: Evidence from the US Theft Law 211

5.1 Introduction 211
5.2 Background 215
5.3 Hypothesized Effects 216
5.4 Data 218
5.5 Analysis of Crime Reporting 220
5.6 Analysis of Raising the Felony Theft Thresholds 226
5.7 Conclusion 233

CONCLUSION 262

LIST OF TABLES

1 Summary Statistics 31
2 Baseline Characteristics 33
3 RD Estimates of the Effect of Female Mayors on Homicides 35
4 RD Estimates of the Effect of Female Mayors on Pre-Existing Homicides 37
5 Placebo Tests Using Alternative Thresholds 37
6 RD Estimates of the Effect of Female Mayors on Homicides by Year . 38
7 RD Estimates of the Effect of Female Mayors on Homicides in the Subsequent Mayor Term 39
8 RD Estimates of the Effect of Female Mayors on Probability of at Least One Homicide Occurring 42
9 RD Estimates of the Effect of Female Governors on Homicides 43
10 RD Estimates of the Effect of Female Mayors on Domestic Homicides 44
11 RD Estimates of the Effect of Democratic Mayors on Homicides against Women 45
12 RD Estimates of the Effect of Female Mayors on Police Responsiveness 46
13 RD Estimates of the Effect of Female Mayors on Crime Reporting . . 47
14 Placebo Tests Using Alternative Thresholds 48
A1.1 Summary Statistics 49
A1.2 Robustness of Estimates to Varying Bandwidths 50
A1.3 RD Estimates of the Effect of Female Mayors on Homicides Involving Female Victims Only: A Quadratic Polynomial 51
A1.4 RD Estimates of the Effect of Female Mayors on Homicides Involving Female and Male Victims 52

A1.5 RD Estimates of the Effect of Female Mayors on Homicides (Excl. Election Years) 53
A1.6 Baseline Characteristics 54
A1.7 Robustness of Estimates to Varying Bandwidths 55
A1.8 RD Estimates of the Effect of Female Mayors: A Quadratic Polynomial 56

1 Summary Statistics 82
2 OLS Estimates of the Appointment of Female Chief Officers on Sexual Offences and Rape Offences 83
3 OLS Estimates of the Appointment of Female Chief Officers on Homicides against Women 84
4 OLS Estimates of the Appointment of Female Chief Officers on Transferred/Cancelled Sexual Offences and Rape Offences 85
5 OLS Estimates of the Appointment of Female Chief Officers on Other Crime Types 86
6 Ex-Post Minimum Detectable Effect Sizes (MDE) 87
A2.1 OLS Estimates of the Appointment of Female Chief Officers on Homicides against Women: Robustness 93
A2.2 OLS Estimates of the Appointment of Female Chief Officers on Crime (excl. London MPS) 94

1 Balance Table 128
2 Estimates of the Effect of Statues (Short Term) on memory stimulus and aspirations 130
3 Estimates of the Effect of Statues on Mean Score of Math and Computer (One/Six-Month) 131
4 Estimates of the Effect of Statues on Language and Social Science (One/Six-Month) 132
5 Heterogeneity in Treatment Effect by Gender 133
A3.1 Summary Statistics 138
A3.2 Estimates of the Effect of Statues by Subject (One/Six-Month) 139

A3.3 Estimates of the Effect of Statues (Short Term) 140
A3.4 Gender Heterogeneity in Treatment Effect on Exam Score by Subject 141
A3.5 Attrition Table 142
A3.6 Attrition Regression 142
A3.7 Multiple Hypothesis Tests for Estimates of the Effect of Statues (One/Six-Month) 143
A3.8 Sensitivity Test for Excluding Outliers of Survey Duration (One/Six-Month) 144

1 State Medicaid Expansion 179
2 Summary Statistics of Key Variables 181
3 Summary Statistics 182
4 OLS Estimates of the Effect of Medical Care Price on Domestic Homicide Rate: Female Victims 183
5 OLS Estimates of the Effect of Medical Care Price on Domestic Homicide Rate: Male Victims 184
6 OLS Estimates of the Effect of Women's Private Insurance Coverage on Domestic Homicides 185
7 Descriptive Statistics of Premium Insurance Tax 185
8 Estimates from the Horse Race Analysis 186
9 Estimates of the Effect of Insurance Coverage on Female Domestic Homicides 187
10 Estimates of the Effect of Insurance Coverage on Male Domestic Homicides 188
11 Estimates of the Effect of Insurance Coverage on Homicides against Women 189
12 Estimates of the Effect of ACA Medicaid Expansion on Domestic Homicides against Women 190
13 Estimates of the Effect of ACA Medicaid Expansion on Domestic Homicides: Female Victims Aged 18 and above 191

14 Estimates of the Effect of ACA Medicaid Expansion on Domestic Homicides: Female Victims Aged under 18 192
A4.1 Estimates of the Effect of Premium Tax on Public Insurance Coverage 193
A4.2 Estimates of the Effect of Premium Tax on Women's Economic Welfare 206
A4.3 Distribution of Primary Source of Payment for Medical Care Resulting from Domestic Violence against Women 206
A4.4 Estimates of the Effect of ACA Medicaid Expansion on Domestic Homicides against Men 207

1 Balance Table 235
2 RD Estimates of the Effect of Felony Threshold on Crime Reporting . 238
3 RD Estimates of the Effect of Theft Felony Threshold on Crime Reporting for Varying Bandwidths 239
4 Donut RD Estimates of the Effect of Theft Felony Threshold on Crime Reporting 240
5 RD Estimates of the Effect of Felony Threshold on Crime Reporting for Rounding Controls 241
6 RD Estimates of the Effect of Theft Felony Threshold on Reporting of Other Property Crime 242
7 Placebo Tests Using Alternative Thresholds 243
8 States Raising the Felony Theft Thresholds (2001-2016) 244
A5.1 RD Estimates of the Effect of Theft Felony Threshold on Crime Reporting for Alternative Polynomials 250

LIST OF FIGURES

1 Mixed-Gender Election Cities 32
2 Homicides and Female Margin of Victory 34
3 Homicides and Female Margin of Victory 36
4 Police Responsiveness and Female Margin of Victory 40
5 Crime Reporting and Female Margin of Victory 41

1 Police Forces in England and Wales 80
2 Female Chief Appointment and Crime 81
3 Event Study Analysis for Sexual Offences 88
4 Event Study Analysis for Rape Offences 89
5 Event Study Analysis for Homicides against Women 90
6 Event Study Analysis for Sexual Offences (Female Chiefs Who Stay) 91
7 Event Study Analysis for Rape Offences (Female Chiefs Who Stay) . 91
8 Event Study Analysis for Homicides against Women (Female Chiefs Who Stay) 92

1 Timeline 127
2 Experiment Assignment 128
3 Conceptual Framework 129
4 Mahatma Gandhi 134
5 Indira Gandhi 135
6 Arjuna 136
7 B.R. Ambedkar 137

1 Trends in Domestic Homicides 178

2 Domestic Homicide and Medical Care Price 180
3 Female Domestic Homicides and Insurance Coverage 183
4 Insurance Premium Tax and Insurance Coverage 184
5 Status of State Medicaid Expansion Decisions 186
6 Event Study Analysis for Domestic Homicides 193
7 Event Study Analysis for Domestic Homicides: Female Victims Aged 18 and Over 194
8 Event Study Analysis for Domestic Homicides: Female Victims Aged Under 18 195
9 Event Study Analysis for Domestic Homicides: Extensive Margin . . 196
10 Distribution of Placebo Estimates 197
11 Event Study Analysis for the Effect of Medicaid Expansion on Women's Dependent Coverage 198
12 Event Study Analysis for the Effect of Medicaid Expansion on Female Unemployment 199
13 Event Study Analysis for the Effect of Medicaid Expansion on Domestic Homicides: Above Median Decrease in Dependent Coverage . 200
14 Event Study Analysis for the Effect of Medicaid Expansion on Domestic Homicides: Below Median Decrease in Dependent Coverage . 201
15 Heterogeneity by Policing 202
16 Event Study Analysis for the Effect of Medicaid Expansion on Prevalence of Mental Illness 203
17 Event Study Analysis for the Effect of Medicaid Expansion on Prevalence of Serious Mental Illness 204
18 Event Study Analysis for the Effect of Medicaid Expansion on Non-Domestic Homicides against Women 205
19 Event Study Analysis for Domestic Homicides 208
20 Event Study Analysis for Domestic Homicides: Female Victims Aged 18 and Over 209

21 Event Study Analysis for Domestic Homicides: Female Victims Aged Under 18 210

1 Crime Situation 234
2 Distribution of Theft Felony Thresholds 234
3 Density of Cash Theft Value 236
4 Effect of Felony Threshold on Crime Reporting 237
5 Event Study Analysis for Theft Clearance Rate 245
6 Event Study Analysis for Police Arresting 246
7 Event Study Analysis for Theft Incidence 247
8 Event Study Analysis for "Home Insurance" Searches 248
9 Event Study Analysis for "Phone Insurance" Searches 249
10 Appendix Figure 1: Balance of Characteristics 251
11 Appendix Figure 1: Balance of Characteristics (Continued) 252
12 Appendix Figure 1: Balance of Characteristics (Continued) 253
13 Appendix Figure 1: Balance of Characteristics (Continued) 254
14 Appendix Figure 1: Balance of Characteristics (Continued) 255
15 Appendix Figure 1: Balance of Characteristics (Continued) 256
16 Distribution of Theft Felony Thresholds 257
17 Event Study Analysis for Theft Clearance Rate (TWFE DD) 258
18 Event Study Analysis for Theft Clearance Rate 259
19 Event Study Analysis for Arrests of Other Crimes 260
20 Event Study Analysis for Other Crime Incidence 261

ACKNOWLEDGEMENTS

First and foremost, I WISH TO THANK my supervisors Prof. Sayantan Ghosal and Dr. Theodore Koutmeridis for their constant support in navigating the personal and professional challenges of my time as a PhD student. They have given me a great deal of their time in the process of building my own research identity.

Prof. Ghosal has changed my beliefs about the importance of interdisciplinarity and working hard to achieve my research goal. His patience and encouragement were invaluable and made this thesis possible. Dr. Koutmeridis opened up many opportunities for me by giving me an opportunity to work as a research assistant on one of his projects, which crucially helped me develop skills. His guidance and insights as I pursued these research projects were beyond measure.

I am also tremendously grateful to the faculty in the Department of Economics, who organised regular economics seminars that provided great opportunities for young researchers like me to speak to and learn from outstanding economicsts from all over the world. I am thankful for the faculty who allowed me to turn to them with all aspects of academia. I feel lucky I got to do my PhD in such a collegiate environment.

I wish to express my gratitude to Prof. Sonia Laszlo and Prof. Franque Grimard for their detailed comments that have improved my work. I have benefited from discussions with Prof. Sonia Bhalotra, Prof. Sharun Mukand, Dr. Geethanjali Selvaretnam, as well as, by participants at the 2021 SES Conference, the 1st Discrimination and Diversity Workshop, the 2021 IAFFE Conference, the 4th IZA Workshop, and the McGill CIREQ Lunch Seminar.

Additionally, my fellow PhD students deserved a special credit. The Reading

Group has been very helpful for exchanging research ideas and collecting feedback. I will not forget the sense of community we have. I need to thank Prateek, Adhiraj, Damiano, Mumin, Hualin, Rohan, Fivos, Ruochen, Kaizhao, Max, Ashraful, Jerome, Moritz, Putthi and all the other I might have forgot to mention. I would particularly like to thank Prateek Bhan for keeping sharing hope with me. His kindness and generosity have influenced me beyond my research journey.

Last but not least, I wish to thank my hometown friends and family for their support and unwavering belief in me.

DECLARATION

I declare that the material contained in this thesis has not been used or published before. This thesis is my own work and it has not been submitted for another degree or at another university.

Signed: Jinglin Wen

The third chapter of Jinglin's PhD thesis is based on the working paper "Role models among us: experimental evidence on inspirations and gender disparities set in stones", which I coauthored with Jinglin.

I confirm that Jinglin, as a main contributor of the paper and in accordance with me, received authorisation to include the results as part of his dissertation.

Signed: Prateek Chandra Bhan

INTRODUCTION

The natural distribution is neither just nor unjust; nor is it unjust that persons are born into society at some particular position. These are simply natural facts. What is just and unjust is the way that institutions deal with these facts.

John Rawls,
A Theory of Justice (1971).

Overview This thesis is composed of five chapters that contribute to important issues in economics and their applications in the context of gender, justice and education. The first two chapters examine whether female leadership in local government and law enforcement agencies reduces violence against women. The third chapter examines the effect of statues on children's development. The fourth chapter examines whether improving access to health care decreases domestic violence. The final chapter examines whether a jurisdiction's sanctions regime influences the decision of victims to seek justice.

Chapter One The first chapter of my thesis examines the effects of electing female mayors on violence against women, using a regression discontinuity design. This chapter makes two main contributions. First, I add to the seminal work on female leadership and policy preferences (Chattopadhyay and Duflo, 2004) and show that elections of female mayors decrease violence specifically targeted toward women but leave other violence unaffected.

My second contribution is to examine the potential mechanisms underlying the relationship. I find strong evidence for the deterrence and reporting channel: police are more likely to arrest the offenders of violence against women and female victims are more likely to seek police involvement. These findings accord well with those documented in Iyer et al. (2012) in the context of India.

Chapter Two The second chapter of my thesis focuses on female leadership in law enforcement agencies, one of the most male-dominated public service fields and examine the effects of the appointment of female chief officers on violence against women. This chapter makes contributions to the literature, because this is the first paper exploring the effects of chief policewomen, which diverges from existing literature with a focus on the distinctive advantage of policewomen in direct interactions with female victims (Miller and Segal, 2019; Amaral et al., 2021; Kavanaugh et al., 2018; Perova and Reynolds, 2017).

Chapter Three The third chapter of my thesis investigates whether installations of statues influence child development in India. This chapter makes contributions to the literature, because it contributes to a growing literature on role-modelling effects from movies and video clips on education attainment (Riley, 2022). We find that a virtual tour to statues leads to an improvement in students' academic performance.

Chapter Four The fourth chapter of my thesis studies whether health care matters for domestic violence. This chapter makes two contributions to the literature. First, it contributes to a growing literature on the impact of healthcare coverage on crime and is the first to study the impact on domestic violence. I find that improved access to health care decreases domestic violence. The results for different measures of accessing health care are quite consistent, speaking to the strength of the results.

Second, exploration of mechanisms suggests that the Medicaid expansion can lower domestic violence by reducing women's economic dependency on their spouses. These findings echo those in the literature documenting that an increase in women's economic independence by improving female labor market conditions decreases domestic violence (Aizer, 2010; Anderberg et al., 2016; Tur-Prats, 2019). I consider alternative explanations and find that the observed patterns are unlikely to be explained by deterrence and mental health accounts.

Chapter Five The fifth chapter of my thesis investigate the effects of the sanction severity on the behavior of victims and criminals. This chapter makes two contributions to the literature. First, it fills the research gap on the relationship between the sanction regime and crime reporting. This is the first study offering causal evidence that the threshold rule causes thefts with value just above the felony threshold to be around 14 percentage points more likely to be reported. This estimate is economically important, when compared to mean reporting of 14 percentage points.

Second, leveraging a dynamic difference-in-differences design, I find that raising theft felony thresholds leads to an increase in theft incidents. In this regard, my work contributes to the literature on deterrence that refers to the behavioral reduction in crime due to offender anticipation of punishment (Becker, 1968).

Organization of the Thesis The next five chapters consist of the essays just described, which are presented as self-contained journal-style articles. They all include their own sections on introduction, data description, identification strategy, results, and discussion. My contributions to existing research are outlined within each chapter.

CHAPTER 1

Female Mayors and Violence against Women: Evidence from the U.S.

Abstract: This study examines the effect of female leadership in local government on violence against women. Regression discontinuity estimates show that elections of female mayors decrease violence specifically targeted toward women but leave other violence unaffected. The effect is strong only for local elected leaders and persistent throughout their term, while there is no effect for non-local leaders. Moreover, the study explores behavioral responses by victims. Evidence suggests that female victims are more likely to report violence against them after female mayors take office. Importantly, female victories are followed by greater police responsiveness to violence against women. There are no such effects for violence against men. These findings survive various robustness checks.

1.1 Introduction

In the United States, more than one in three women experience sexual violence involving physical contact during their lifetimes (Centers for Disease Control, 2018b). Rapes result in nearly $3.1 trillion to the economy over the lifetimes of all 25 million victims (Peterson et al., 2017). Importantly, the economic burden would be much larger after accounting for sexual violence excluding rapes. Empowering women in political decision making to eliminate violence against women has been identified as a priority in many countries (United Nations, 2012).

Does female leadership decrease violence against women? On one hand, female leaders, when compared with male leaders, share common concerns about being victimized, therefore, they are more likely to place women's concerns on the political agenda and combat violence against women. For example, mayor of Atlanta, Shirley Franklin launched the "Dear John" campaign aimed towards reducing sex trafficking, which got the police to take it more seriously.[1] As Franklin herself explained, "As the first woman mayor, I needed address these issues" (Majic, 2017).

On the other hand, it is commonly argued that gender identity norms could create an aversion to women occupying leadership positions. As in Akerlof and Kranton (2000), one person's actions can have meaning for and evoke responses in others if deviating from specific behavioral prescriptions. Because empowering women politically might violate prescribed expectations for women, there could be backlash against women in response to their greater political empowerment. For instance, Dube and Harish (2020) find that single queens are more likely to be attacked by others than single kings.[2]

U.S. police departments are organized at the city level and ultimate decision-making authority on police issues resides with the mayor (Levitt, 1997).[3] Accordingly, the preferences of mayors play a significant role in policing and crime. However, there is little conclusive evidence of whether cities vary in experiencing crime under female versus male leadership. In contrast to this association, there is a dense literature that focuses on the association between female leadership and policy preferences regarding economic development (Chattopadhyay and Duflo, 2004; Clots-Figueras, 2012; Brollo and Troiano, 2016).[4] Earlier work by Iyer et al. (2012)

[1]In North America, the vast majority of detected victims of trafficking continue to be women, reaching nearly 85 percent in 2016 (United Nations Office on Drugs and Crime, 2018).

[2]The literature also sheds light on the relationship between female leadership and conflict on the modern era (Caprioli and Boyer, 2001; Regan and Paskeviciute, 2003; Koch and Fulton, 2011).

[3]There are only a few exceptions (Levitt, 1997).

[4]Political theories of representative bureaucracy also suggest that minority bureaucrats often implement policies to reduce the disparate treatment that minority clients have received historically (Meier and Stewart, 1992). For example, Keiser et al. (2002) take into account the contextual environment that influences whether sex will lead to active representation in the bureaucracy on this topic and find that increasing female representation in bureaucratic agencies helps women in the

does examine the effect of local female leaders on the occurrence of violence against women but the estimate is statistically insignificant in the Indian context where local governments do not have direct power over policing.

In this work, I examine how female mayors affect violence, and in particular violence against women, exploiting the outcomes of mixed gender close electoral races to implement a regression discontinuity (RD) design. While cities where female candidates win and lose by wide margins tend to be different, the results of validity tests show that the outcomes of close elections are uncorrelated with pre-existing city characteristics and violence trends.

To conduct my analysis, I construct a panel data set by combining detailed data sources on crime, elections, and various city-level socioeconomic characteristics. RD estimates document that female victories in mayoral elections decrease homicides involving only female victims by approximately 1.79 homicides per 100,000 population. Importantly, these estimates are economically meaningful, when compared to mean value of 2.59 homicides per 100,000 population.

A reasonable concern with my analysis is that electing a female mayor may have generalized effects on violence, regardless of victim gender. If so, different potential channels other than gender identity of mayors might be in operation. To address this concern, I perform falsification checks by regressing homicides involving male victims on female mayors. The results suggest that having a female mayor does not influence violence against men, increasing my confidence in the results.

The natural question to ask is whether the effect differs over time across the mayor's term of office. That is why I conduct heterogeneity analysis by introducing interaction terms. The corresponding results show that the gender effect persists during the mayor's term of office. In addition, I also examine whether it outlasts her tenure and find no effect after a male succeeds a female as mayor. It implies that temporary exposure to female mayors may not have long-run effects on attitudes and behaviors of potential offenders.

I examine two potential mechanisms in which the presence of female mayors

population.

could lead to a decrease in violence against women. The first mechanism relates to police responsiveness to violence against women. To alleviate women's concerns, female mayors could devise the police and crime plan where fighting violence against women is made a priority. In turn, police may have to take women's grievances more seriously and become more responsive to violence against women. When potential offenders observe an increase in police efforts to protect women, they are less likely to commit crimes against women.

The second mechanism builds on the critical role of crime reporting. Crime reporting has been viewed as an important measure of the performance and efficiency of law enforcement services (Soares, 2004a).[5] Underreporting of violence against women is expected to be particularly worrisome. If the police consider violence against women more serious than they did before elections of female mayors, victims would adjust their behavior accordingly and their propensity to report crimes would increase.

In addition, as in Scrase (2002), women's victories in mayoral elections might lead women to question their social position and might decrease the acceptability of being badly treated. Similarly, they would be more inclined to come forward and report crimes. Given the higher probability of incurring police involvement, potential offenders would be deterred from committing crimes against women in the first place.

To test these mechanisms, I match mixed-gender elections to survey responses on crime victimization. I find that female victories raise the probability that perpetrators are apprehended for having committed crimes against women by 14 percentage points. These results are supportive of the police responsiveness mechanism: the advancement of women in political decision making brings about an increase in police efforts to address violence against women.

Also, I observe a behavioral response of female victims to the electoral success of female candidates. Specifically, female victories increase the likelihood of fe-

[5]Despite its importance, there is limited literature in the economics of crime focusing on crime reporting as an outcome of interest.

male victims reporting violent crime to the police by around 33 percent. This result is consistent with the crime reporting mechanism. Furthermore, the results from falsification checks in particular, are worth mentioning. The RD estimates of the effects of having female mayors on crime reporting by male victims and police quality for violence against men are statistically insignificant, supporting the women-specific channel. The combination of these results accords with deterrence that refers to the behavioral reduction in crime due to offender anticipation of punishment (Becker, 1968; Levitt, 1998; Chalfin and McCrary, 2018).

This study fits into four main strands in the literature. First, it contributes to the economics of crime literature on violence against women by investigating the role of gender identity of mayors in deterring offenders from committing those crimes. The existing literature considers the impacts of gender pay gaps (Aizer, 2010), labor market conditions (Anderberg et al., 2016; Bhalotra et al., 2020), and family types (Tur-Prats, 2019). Studies also link violence against women with emotional cues (Card and Dahl, 2011), mandatory arrest laws (Iyengar, 2009), intergenerational transmissions (Pollak, 2004), and ancestral features (Alesina et al., 2016). Unlike that literature, which mainly focuses on the economic and historical causes of violence against women at the individual level, this study contributes to the literature by primarily examining whether women's political empowerment at the administrative level can help reduce violence against women.

Second, this paper also contributes to the literature on the effect of policing on crime by providing evidence on deterrence. Beginning with Levitt (1997), much of that literature focuses on police staffing, suggesting that the size of the police force significantly reduces crimes (Evans and Owens, 2007; Chalfin and McCrary, 2018). Leveraging a redeployment of police in response to a perceived terrorist threat, studies also investigate the response of crime to changes in normal routines of policing (Klick and Tabarrok, 2005; Draca et al., 2011). Some studies have also focused on the effect of police demographic characteristics on crime (Antonovics and Knight, 2009; Amaral et al., 2021; Perova and Reynolds, 2017; Kavanaugh et al., 2018; Miller and Segal, 2019). That literature concentrates on the management and staffing decision within the law enforcement department, while this study contributes to the literature by exploring the change in governmental leaders due to direct elections. In particularly, directly elected leaders are held accountable for their actions by communities they serve.

Third, this paper also contributes to the literature that relates the identity of leaders to policy outcomes. Many studies have examined the effects of particular types of leader identity, along dimensions such as race (Meier and Stewart, 1992), caste (Pande, 2003), party (Blais et al., 1993; Ferreira and Gyourko, 2009), and gender (Chattopadhyay and Duflo, 2004; Clots-Figueras, 2012; Iyer et al., 2012; Ferreira and Gyourko, 2014; Brollo and Troiano, 2016). My study builds on this literature by showing how the gender identity of mayors could be consequential for the behavior of police, victims, and criminals relating to crime situations. Another strand of related studies has also suggested that the presence of women in corporate leadership affect firm outcomes (Ahern and Dittmar, 2012; Matsa and Miller, 2013; Bertrand et al., 2019). This paper also looks to the impact of identity of political leaders, but instead of focusing on economic outcomes, I focus on the outcome of public safety, particularly for women, and argue that women as leaders can help and protect women from violence in the population.

Fourth and finally, I view my work as closely related to quantitative and qualitative studies of underreporting of violence against women. The main barriers to reporting include stigma that will tarnish the woman's own reputation (Kishor and Johnson, 2005; Overstreet and Quinn, 2013), cultural beliefs (Njuki et al., 2012), economic dependence (Wolf et al., 2003), fear of retaliation (World Health Organization, 2005; Kishor and Johnson, 2005), and discriminatory and stereotypical attitudes toward victims in law enforcement settings (Allen, 2007; Maier, 2008; Belknap, 2010; Palermo et al., 2014). This study contributes to the literature by demonstrating the critical role of female leadership in crime reporting and access to justice. My goal in this paper is to examine whether creating female-friendly political environment can encourage more female victims to come forward to the criminal justice system and bring justice for them.

The remainder of the article proceeds as follows. Section 2 presents mechanisms through which female leadership can influence violence against women. Section 3 describes the data. Section 4 conducts the analysis of female mayors and homicides. Section 5 does the same for police responsiveness and crime reporting. In Section 6 I discuss my results. Section 7 concludes.

1.2 Mechanisms

1.2.1 Police Responsiveness

One account of how female leadership can influence violence against women focuses on police responsiveness. As U.S. mayors have direct power over local police plans and budgets, women taking office could put the achievement of gender justice on the policing agenda and even prioritize the prevention of violence against women (analogous to the effect of women as policy makers on the expansion of infrastructure that is directly relevant to the needs of their own gender in Chattopadhyay and Duflo (2004)).

If law enforcement departments allocate more police resources to investigate crimes against women such as follow-up home visits, and apprehend offenders, potential offenders would expect higher costs from committing those crimes. As a result, female leadership in local government can lead to a rise in police responsiveness to crimes against women and in turn deter those crimes.

1.2.2 Crime Reporting

A second account of female governance and violence against women leans on the importance of crime reporting in triggering a response by the criminal justice system. There is a dense literature that documents that poor treatment by the police, often referred to as "re-victimizing" is one of the major barriers for female victims to reporting (Allen, 2007; Maier, 2008; Belknap, 2010; Palermo et al., 2014). For example, one study shows that half to three fourths of all police reported sexual assaults are filtered out of the legal system at the police level (Du Mont and Myhr, 2000).

If female leadership results in greater investigative efforts by police, then the benefits to victims of reporting will increase. That is to say, given the risk of social recrimination and privacy loss linked with reporting, higher probability of apprehension, implying availability of justice, increases victims' incentive to report. Not surprisingly, reporting by victims would reinforce the probability of detection and the expected cost of violence against women. In short, crime reporting can function as a powerful disincentive for potential offenders to commit crimes.

Furthermore, the acceptability of beating is also an important reason why a great number of female victims decide not to report those crimes. Although the law denies men the privilege of beating their wives (Siegel, 1995), World Values Survey data for the U.S. show that a small number of women continue to justify wife-beating as a husband's right.[6] Probably, female victories in mayoral elections expose women to strong women characteristics that are radically different from their own behaviors and family values, and they start to question their status. Thus, female leadership might generate a "role model" effect that can raise awareness of women's rights and decrease women's tolerance for violence (analogous to the effects of cable television on lifestyles, behaviors, and, in particular, the acceptability of domestic violence toward women in Jensen and Oster (2009)).

1.2.3 Backlash

A third account of how women occupying leadership positions might influence violence against women relates to the retaliation behavior. As political institutions remain heavily male-dominated,[7] an increase in women's political power through electing female mayors may trigger backlash that takes form of violence from those with traditional gender stereotypes. This is consistent with recent work finding that

[6] https://www.worldvaluessurvey.org/WVSOnline.jsp

[7] Women are still underrepresented in political institutions in the U.S. (Center for American Women and Politics, 2017; Lawless and Pearson, 2008).

unmarried queens are attacked more than kings (Dube and Harish, 2020). In addition, this backlash account is shared with studies that find an increase in women's bargaining power through empowering them economically leads to backlash from their spouses, who may prefer that women do not work (Field et al., 2016).[8]

1.2.4 Empirical Implications

The accounts above result in the following empirical implications. If the police responsiveness account holds, having a female mayor would decrease violence against women. Importantly, if the crime reporting mechanism is also at work, the crime-reduction effect is expected to be especially large. In contrast, if backlash account holds, having a female mayor should bring about an increase in violence against women.

1.3 Data

This section describes the sources of information for my data. My analysis matches data on mayoral elections to crime, with the aim of studying the behavior of offenders, police, and victims after elections of female mayors.

1.3.1 Data on Homicides

The data source for homicides is the Supplemental Homicide Reports (SHR) within the Uniform Crime Reporting (UCR) program. Underreporting of crime is an important concern that results in biased estimates, while it is likely to be the least for homicides. That is why violence is measured by homicides in this study. The SHR provides detailed information about the victim and the circumstances of the murder. Because my analysis focuses on the gender effect, I use information on

[8] Other backlash studies suggest that empowering women economically establishes the incentives of men to use violence or threats of violence as a way of extracting rents from their female partners (Eswaran and Malhotra, 2011; Bobonis et al., 2013; Field et al., 2016; Erten and Keskin, 2018)

gender of victims. In doing so, I am able to examine the effects of female mayors on murders targeted toward women and men, respectively.

Specifically, in the main analysis I categorize homicides into two types: (i) homicides involving female victims only, (ii) homicides involving male victims. As a robustness check, I also use data on homicides involving female and male victims. The main outcome variables of interest are the city-year homicide rates by victim sex, calculated by dividing the homicide counts respectively by the population (in 100,000s) in the city. Panel A of Table 1 displays summary statistics of homicides by victim sex in the sample used for the main analysis. In the same vein, I employ the homicide data to construct measures of the violence extensive margin as alternative outcome variables. Additionally, to examine the effect of female governors on homicides, I collect the state-level homicide data and compute homicide rates by victim sex.

1.3.2 Data on Police Responsiveness and Crime Reporting

Data on police action and crime reporting are obtained from the Bureau of Justice Statistics' (BJS) National Crime Victimization Survey (NCVS), which is the nation's primary source of information on criminal victimization. To be more specific, I use the extract file created from the NCVS for a sample of Metropolitan Statistical Areas (MSAs) from 1970s through 2000s. This survey contains select household, person, and incident variables for persons who reported a crime.

As the nation's primary source of information on criminal victimisation, uniform definitions of victimisation are applied in the National Crime Victimization Survey (NCVS). Particularly, the determination of whether an event is a crime and if so, the type of crime, is not decided during the interview. Rather, the characteristics of each reported event are detailed during the interview. Final classification of the crime type occurs during data processing and is based on information provided in the incident report (Lauritsen and Catalano, 2005). Thus, there are no definitional variations across cities. Since 1973, the NCVS has started to measure the extent of victimization by the major crimes commonly referred to as rape, robbery, assault, burglary, motor vehicle theft, and theft. In 1992, the NCVS added new crimes, yet the definitions of existing crimes did not change (Federal Bureau of Investigation, 2019).

The two main outcome variables are used to measure the behavior of police and victims, respectively. The first is an indicator variable of whether police arrested the offender, serving as a measure of police responsiveness. The second is an indicator variable of whether or not the crime was reported to the police. This survey also includes information at the individual level such as the respondent's sex, race, Hispanic origin, marital status, and age group.

Typically, each MSA is centered on a single large city (Office of Management and Budget, 2010). I construct the dataset of the crime-female mayors to examine whether police and victims respond to elections of female mayors in the principal city. After 1993, the NCVS adds crimes not previously measured. To be consistent, my analysis focuses on crimes measure throughout the sample period. Crimes in this survey include both violent crime and property crime, but the analysis sample is restricted to violent crime. Panels A and B of Appendix Table A1.1 display summary statistics of the two main outcome variables.

1.3.3 Data on Elections

The data on mayoral elections are obtained from multiple sources, including Internet searches[9], Freedom of Information Act (FOIA) requests and Penn Institute for Urban Research. Information includes the name of the winner and the runner-up, gender of candidates, vote totals for each of two candidates, and total votes counted, party affiliation, and the timing of elections between 1950 to 2007.

As the homicide data from SHR available start in 1976, I restrict my sample period from 1976 to 2007. Table 1 offers summary statistics of female participation and performance. The numbers of total elections and mixed-gender elections

[9]State, county, and city government and legislature archive websites and Ourcampaigns

where mayors and second-place candidates are of different gender are reported in Column 1 of Panel B, respectively. Over the period, there are 812 mixed-gender elections matched to homicide data, which comprise approximately 20 percent of elections in the dataset (see Column 2 of Panel B).[10]

Panel C reports the statistics for vote share by candidate sex in mixed-gender elections. The average vote shares for female and male candidates are 0.48 and 0.47, respectively. The average female margin of victory is 0.01, suggesting that female and male candidates have an almost equal chance of winning. This ties into the literature on female political participation that shows the gender of a candidate has little impact on a candidate's probability of winning (Anastasopoulos, 2016). The mixed-gender election cities are shown in Figure 1. My analysis also matches 86 mixed-gender elections to survey responses on crime victimization from the NCVS. The corresponding summary statistics are displayed in Panels C and D of Appendix Table A1.1.

For gubernatorial elections, I collect data from the Candidate and Constituency Statistics of Elections in the United States, the VoteSmart Biographical database, and archives of state and legislature. In order to match gubernatorial elections to homicide data, I end the sample period in 2018. Over the period of my investigation, there are 114 elections where governors and second-place candidates are of different sex matched to the homicide data.

1.3.4 Data on Characteristics

Data on police employment information come from the Law Enforcement Officers Killed or Assaulted (LEOKA) collection. It includes counts of police employees of each sex at the city level. To take account of the size and payroll of city government, I also collect data on the number of public employees and payroll per public employee from the Annual Survey of Public Employment and Payroll (ASPEP).

Information on geographic characteristics such as land area and elevation is

[10]As information is only available for the winner and the runner-up, a mixed-gender election is defined as a race where the winner and the runner-up are of different gender in this study.

from U.S. Census and U.S. Geological Survey. I supplement my analysis with additional data on city social-economic characteristics including income per capita and unemployment rate from the U.S. Census. Since they are available in census years, I linearly interpolate between Census years. For data on social-economic characteristics at the MSA level and state level, I obtain annual measures from the Current Population Survey (CPS) and U.S. Bureau of Economic Analysis (BEA), respectively. Population data for each analysis are obtained from the UCR.

1.4 Effects of Female Mayors on Homicides

This section focuses primarily on whether female victories in mayoral elections have significant impacts on homicides, providing direct evidence that offenders respond to female leadership. It first discusses the empirical strategy. Then it performs the regression analysis and presents the main results. In the last part of the section, I conduct several robustness tests.

1.4.1 Empirical Strategy

I begin with the homicide rate because it is a measure of the most severe form of violence and a major policy concern. More importantly, issues of endogenous reporting should be of much less importance for homicides than other crime types. This section discusses the implementation of the regression discontinuity (RD) design used to gauge the impact of female mayors on homicides. The RD approach exploits a discontinuity in the treatment assignment to identify causality (Imbens and Lemieux, 2008; Lee, 2008). That is, the gender of a mayor changes discontinuously at the threshold between a female victory and loss.

The main endogeneity issue is that the assignment of mayor gender correlates with city characteristics, particularly in a city where female wins or loses by a large margin. However, under the assumption of continuity of city characteristics at the treatment threshold, the RD estimator identifies the treatment effect at the specific cutoff value of the female win margin. Before examining the plausibility of the design's identification assumptions, I specify the following regression model for estimating the RD treatment effect:

$$\begin{aligned} Y_{at} = \alpha_0 + \alpha_1 FemaleMayor_{aT} + \alpha_2 FemaleMayor_{aT} \times VoteSpread_{aT} \\ + \alpha_3 VoteSpread_{aT} + \alpha_4 X_{at} + \sigma_a + \phi_t + \varepsilon_{at} \end{aligned} \tag{1}$$

Where Y_{at} is the outcome of interest in city a in year t. $FemaleMayor_{aT}$ is a binary indicator that equals one in city a, where, in the most recent election year T, a female is elected as mayor. $VoteSpread_{aT}$ is the margin of female victory, defined as the difference between the percentage of votes received by the female candidate and the percentage of votes received by the male candidate.[11]. X_{at} represents a set of controls capturing city characteristics such as female officer share, total police forces, unemployment, income per capita, and population. Adding controls is not necessary for the identification but improves the efficiency of the estimation. σ_a and ϕ_t are city and year fixed effects, respectively, and ε_{at} is the error term.

The estimator should be interpreted as a weighted average treatment effect (ATE) of female mayors. As in Lee and Lemieux (2010), in the context of election, it is appropriate and natural to interpret the effect of treatment "female mayor by virtual of women winning elections whereby positive margin of female victory is required to win" as an ATE weighted by the probability of being close to the threshold of female win or lose, because there are no other thresholds deciding election results.

A mayor's governance lasts for a two-year or four-year term, while the outcome variable of interest is at the city-year level. Hence, α_1 captures the average annual effect of electing a female mayor across the electoral cycle. As the bandwidth selectors commonly generate a "large" bandwidth, giving rise to data-driven confidence intervals that could be biased, I follow the methodology proposed by Calonico et al. (2014) where a "small" enough bandwidth is selected. The optimal bandwidth ac-

[11] As Gelman and Imbens (2019) suggest that high-order polynomials should not be used in RD design, I only report the results of specifications use linear RD polynomials. Nevertheless, the estimates with a quadratic RD polynomial, both in terms of magnitude and significance, are strikingly similar to those with a linear RD polynomial (see Appendix Table A1.3)

cording to the selection procedure is 0.07, which is used in the main analysis. Nevertheless, the results from regressions using the Imbens-Kalyanaraman bandwidth show qualitatively similar patterns.[12]

RD strategy requires that baseline covariates are balanced across the treatment threshold. Table 2 reports the results of balanced covariate checks. Columns 1 and 2 present the mean values for various city baseline characteristics, which relate to population density, labor market and the economy, geography, incumbent gender, policing, and public employee payroll. There are no average differences between cities with female and male mayors, except for female officer share (see Columns 3 and 4). When I use the RD specification to test for discontinuous changes over the cutoff point, I find that in no case are estimates statistically significant (see Columns 5 and 6). These results suggest that there are no discontinuous effects at the threshold for the pre-characteristics, and therefore offer strong support that the identification strategy is valid.

RD strategy also requires that the density of the running variable is balanced across the treatment threshold. If the elections are rigged, the causal interpretation of RD strategy would be undermined. To test such manipulation, I examine the distribution of the running variable (see Figure 2). Panel A first shows a histogram of the female margin of victory for the entire range. Panel B shows the McCrary (2008) test of whether there is a discontinuity for entire range in the density of the female margin of victory; the McCrary test statistic is 0.05 (standard error 0.18). Neither figure reveals any discontinuous change at the female win-loss threshold. Also, I implement the McCrary test for the close election sample, and the estimate is considerably small and statistically insignificant. To sum up, these results show that neither the female nor male candidates systematically wins the elections.

[12] Imbens and Kalyanaraman (2012)

1.4.2 Main Results

I start by presenting treatment estimates on two types of homicides: homicides involving female victims only and involving male victims. Panels A and B of Table 3 display the RD estimates of the impact of female mayors on each outcome, along with standard errors. Columns 1 through 4 report results without controls, with police demographic control, with police demographic control plus total police forces control, and with full controls, respectively.

Through the different specifications, the coefficients in Panel A are statistically significant, strongly implying that the estimated effect is thus quite insensitive to the inclusion of control variables. The magnitudes of the coefficients are close across specifications, ranging from -2.23 without any covariates to -1.79 with all covariates. The preferred specification is displayed in Column 4 with full set of control variables. The estimate suggests a negative and statistically significant effect of female mayors on homicides involving female victims only: electing a female mayor decreases homicides involving only female victims by around 1.79 homicides per 100,000 population.

The absence of a positive impact allows me to interpret these results as evidence that there is no significant backlash against women with respect to women's safety in the population in response to increasing female political empowerment. Nevertheless, it is worth noting that there is no conclusive evidence of whether female leadership can influence overall backlash, as backlash against women could take different forms. For example, Ayalew et al. (2018) present evidence of gender discrimination by subordinates and find that administrative employees are less likely to follow guidance provided by a female leader, using an experimental design.

A way to corroborate my results is to test whether violence against men is also influenced by the presence of female mayors too. This could be a key criticism, as violence against women and female mayors are very closely associated through gender identity, yet the same is not true for violence against men. If I detect a significant impact for violence against men as well, beyond and above violence against women, then different potential channels other than the gender-specific channel could be at work, namely, systematically changes in institutional settings after female mayors take office affect violence in general.

Panel B of Table 3 reports RD estimates, that is, the effect of having female mayors on homicides involving male victims. Across all specifications (see Columns 1 to 4), the coefficients are not statistically significant. A comparison of the significant estimates in Panel A with the insignificant estimates in Panel B suggests that female victories do affect violence that is specific to female victims, as having more sympathetic women leaders could serve women better, while general violence is unchanged. I interpret these results as evidence that there is no overall change in law and order conditions other than female victories in mayoral elections that are driving my results.

Even though the results from RD strategy document a strong and significant treatment effect of female mayors on violence against women, they still leave room for visual inspection. Figure 3 plots the relationship between the homicide measures and the female margin of victory, with a negative margin representing a female loss. The solid line represents predicted values from a regression of residuals from controls and fixed effects on a linear polynomial in the vote spread, estimated separately on either side of the treatment threshold, with the dashed lines identifying the 95% confidence intervals.

I see a clear discontinuity at the female win-loss threshold while the relationship appears to be smooth elsewhere (see Panel A). In contrast, Panel B shows that female victories do not influence homicides involving male victims. The graphs confirm the tables, suggesting a significant effect of female males on violence against women, but little clear effect on violence against men.

Imbens and Lemieux (2008) recommend sensitivity analysis of the RD model with respect to the bandwidth. Appendix Table A1.2 presents the RD estimates of the impact of electing a female mayor on homicides involving female victims only for various bandwidths. The estimated effect is statistically significant across bandwidths. Thus, my results are robust to varying the bandwidth choice. As a further check, I examine the impact of female mayors on homicides involving female and male victims which are not gender-specific. Once more, I find no significant changes in crimes not specifically targeted toward women after elections of female mayors (see Appendix Table A1.4).

There is a concern in relation to the timing of mayoral elections, which builds on the fact that the elections across cities are not held at the same time in election years. I tackle this issue by excluding the sample in election years. Appendix Table A1.5 reports the corresponding results by victim sex in the same way as with the other RD analysis: female victims only and male victims. Both sets of results are consistent with existing results, suggesting that such a timing concern is not posing a threat to my research design.

Overall, the results from the RD strategy with graphical analysis paint a very consistent picture, suggesting that female victories in mayoral elections induce a decrease in homicides where women are specifically targeted because of their gender.

1.4.3 Pre-existing Homicides and Female Victories

The key assumption underlying my identification strategy is that a city where the female barely lost is an appropriate counterfactual for a city where the female barely won. If homicides during the term of the predecessor have a strong relationship with the outcome of mayoral elections, the assumption would be violated. This section therefore examines the validity of my RD strategy by regressing homicides during the pre-election period, that is, the predecessor's term of office.

Specifically, as in the previous exercise, I estimate the RD specifications using homicides involving female victims only and male victims as the dependent variables, respectively. In Table 4, Columns 1 and 2 report estimates for the former without controls and with controls, whereas Columns 3 and 4 report estimates for the latter. Regardless of the gender of victims, the coefficients are statistically identical to zero, implying that election outcomes are not driven by the pre-existing trend in homicides. Hence, the results rule out the possibility of selections in cities and corroborate that what I detect is the impact of female mayors.

1.4.4 Placebo Thresholds

Another test for the validity of my identification strategy is required to tackle the concern relating to noise in the data by placebo tests for different thresholds values for the margin of victory. As suggested by Imbens and Lemieux (2008), these tests are conducted by including the subsamples only to the right or only to the left of the true cutoff value, so as to avoid including in the estimation a point where there is a discontinuity.

Again following Imbens and Lemieux (2008), I split the sample at the true cutoff value of zero and run the RDD estimation at the thresholds equal to the median margin of victory of each subsample to increase the power of the test to find jumps. If the effect observed at zero is a genuine one, then there should be no effect observed at other thresholds. Panel A of Table 5 reports the corresponding results for subsamples of the left and right of the true threshold, respectively, showing neither of estimates is statistically significant. Then I split the sample again at the new median thresholds and set another threshold at the median of the new samples. Panel B of Table 5 displays the corresponding results. At those newly generated four thresholds, there is no significant discontinuity. Therefore, I conclude that there is little evidence that my results are due to noisy jumps in the data.

1.4.5 Persistence Throughout Her Term

One may wonder whether the effect of a female mayor is different by the year of her term, as my design so far is using one dummy for the entire period of the mayor's tenure and masking the heterogeneous effects. To shed light on such effects, I interact the female win indicator, vote spread, and female win indicator $\times$ vote spread with indicators for the years of the mayor's term, respectively.

Table 6 exhibits the RD estimates for specifications with flexible controls. In the first two columns the dependent variable is the homicides involving female victims only, while in the last two columns the dependent variable is the homicides involving male victims. Although there is a slightly falling trend in the magnitude of the RD estimates over time, the estimated effects on homicides against women only for different years of the mayor's term of office are large, positive, and statistically significant. Also, the estimates are very robust to different specifications (see Columns 1 and 2).

Once again, I find no significant results for homicides against men that do provide evidence consistent with the gender-specific mechanism (see Columns 3 and 4). The combination of these results implies that the effect of a female victory remains persistent over the tenure.

1.4.6 Effects in the Subsequent Male Term

So far I have provided evidence suggesting that there is a statistically significant effect of female victories in mayoral elections on violence against women and such an effect remains significant throughout the mayor's term. Does the effect outlast the mayor's term? To answer the question, I examine the impact of electing a female mayor on homicides during the subsequent mayor term.

Table 7 reports the corresponding estimates. The dependent variables are homicide rates by victim sex, and the point in time is the duration of the subsequent mayor term. None of the coefficients is statistically significant at any conventional significance level, suggesting that the gender effect fades away after turnovers. This result is justifiable, as male mayors differ in their crime policy formulation and implementation preferences. For instance, the "Dear John" campaign ended with the term of Mayor Shirley Franklin, since her successor, Mayor Kasim Reed, did not continue it (Majic, 2017).

Nonetheless, this result is important, because it implies that violence against women is quite sensitive to female mayors. In other words, I derive a strong and significant elasticity for violence against women with respect to the gender identity of local leader. Also, this result indicates that temporary special measures such gender quotas, aimed at enhancing female political leadership might not lead to a long-standing change in violence against women.

It is worth noting that the probabilities of winning the elections are very similar for female and male candidates, while female participation is much lower than male participation (see Panels B and C of Table 1). Intuitively, paying more attention to the underlying barriers to women's equal participation in local politics could be essential to ensure sustainable representation, and thus combat violence against women in the long run.

1.4.7 Extensive Margin

An important question to ask is whether female victories bring about a decrease in violence against women only in places that would have experienced some violence against women regardless? What about places where the frequency of violence is relatively low? To offer a better understanding of the full picture, I examine the effect of female victories in mayoral elections on the probability that at least one homicide occurs. The same exercise is repeated for extensive margin and Table 8 displays the RD estimates by the gender of victims.

Columns 1 and 2 consider homicides involving the female victim only, while Columns 3 and 4 consider homicides involving the male victim. The results suggest that female victories significantly decrease the probability of at least one homicide against women occurring by around 48 percentage points, while they do not affect the probability of at least one homicide against men occurring. Once more, I find that the estimated effect is quite stable and precise when I add a battery of controls.

The results provide clear evidence that female victories influence both the extensive margin and intensive margin of violence against women. Therefore, I offer prima facie evidence that women leaders are not just responding to deteriorations in women's safety. Rather, it is the alignment between their own preferences and the preferences of women in the population that generates a powerful impact on violence against women. This finding supports the earlier work by Chattopadhyay and Duflo (2004) suggesting that female political leaders prefer programs more closely linked to women's concerns.

1.4.8 Female Governors and Homicides

Decentralization is widely believed to promise a range of benefits by fragmenting central authority and making government more efficient (Bardhan, 2002). Yet, whether local political autonomy matters for violence against women has been largely underexplored in the literature. In the U.S. context, the law enforcement system is decentralised where law enforcement in each area is in accord with the preferences of local citizens. Notably, the chiefs of police/commissioners of police as the most senior officers in the police departments owe their allegiance to a city, and they are appointed by the mayor of a city (Sampson, 2012). In contrast, governors have limited capacity to supervise local policing action. Do female governors decrease violence against women?

Using the additional data on gubernatorial elections I am able to examine the effect of female governors on violence and tackle the above question. I conduct the same type of RD analysis as in the previous section where the sample is restricted to mix-gender elections and the optimal bandwidth around the threshold is used.[13] Corresponding results using homicides involving female victims only and involving male victims as the dependent variables, are displayed in Columns 1 to 2 and Columns 3 to 4 of Table 9, respectively. None of the results is statistically significant. Thus, increasing female leadership in the state government does not influence local crime patterns.

Comparing results for female mayors with results for female governors is useful and helps to paint a nuanced picture of the effect of governance arrangements on social order. The findings have important implications for criminal justice remedies, given that achieving gender equality at lower levels of governance could be more

[13] Validity checks are also performed and the results provide support for the validity of the RD strategy.

important than that at higher levels of governance.

1.4.9 Other Results

Domestic abuse exists as large part of violence against women. I also examine the effects of elections of female mayors on domestic homicides against women, which include homicides committed by a current or former boyfriend, or spouse of a female victim. Table 10 presents the corresponding results, repeating the previous exercises. Likewise, I find that electing female mayors significantly decrease domestic homicides.

Moreover, one may question whether the effects of female mayors on violence against reflect the effects of their political affiliation. I use data on candidates' party affiliation, examining the effect of electing a Democratic mayor on violence against women. Table 11 shows the RD estimates of Democratic mayors on homicides and domestic homicides against women. None of the coefficients is statistically significant at any conventional significance level, suggesting that the detected gender effect are not driven by mayors' party identity.

I summarize the key results as follows. On the one hand, the results offer strong evidence that there is a genuine decrease in violence against women following female victories in mayoral elections. On the other hand, evidence suggests that female victories do not influence violence against men. Importantly, such effects on violence against women last for the duration of the mayor's term. When it comes to the levels of governance, it is worth noting that women's leadership at the city level could be more influential than that at the state level in the context of violence against women. Overall, my results indicate that maximizing female voice in local politics significantly affects the behavior of potential offenders and reduces gender injustice that has long needed to be tackled.

1.5 Effects on Police Action and Crime Reporting

Having focused on the relationship between female mayors and violence against women, I now investigate whether being exposed to female mayors affects police responsiveness and crime reporting. In doing so, my analysis is able to shed light on the link between female leadership and the behavior of police and victims, uncovering potential mechanisms through which female victories affect violence against women. The section first discusses the identification strategy. It then examines whether female victories induce greater police responsiveness to violence against women. Finally, it explores the impact of female victories on the reporting decisions made by female victims.

1.5.1 Empirical Strategy

To gauge the RD treatment effect of having a female mayor on the behavior of victims and the police, I use an empirical specification of the following form:

$$\begin{aligned} Y_{iat} = \beta_0 &+ \beta_1 FemaleMayor_{aT} + \beta_2 FemaleMayor_{aT} \times VoteSpread_{aT} \\ &+ \beta_3 VoteSpread_{aT} + \beta_4 X_{iat} + \sigma_a + \phi_t + \varepsilon_{iat} \end{aligned} \tag{2}$$

Where Y_{iat} is the outcome of interest for crime i in area (MSA) a and year t. $FemaleMayor_{aT}$ is an indicator variable set to one in area a, if in the most recent election year (T), a female is elected as mayor of its principal city. $VoteSpread_{aT}$ is the margin of female victory.[14] X_{iat} is a set of controls for local and victim characteristics. Local characteristics include female officer share, total police forces, income per capita, and unemployment. Victim controls include race, Hispanic origin, marital status, and age. σ_a and ϕ_t are area and year fixed effects, respectively, and ε_{iat} is the disturbance term. The estimate β_1 captures the effect of female victories on the outcome of interest. The optimal bandwidth according to the method of Calonico et al. (2014) is 0.25.

[14]When using a quadratic polynomial, the estimates remain significant (see Appendix Table A1.8).

Before moving to the RD results, I perform the same validity checks as in the previous section. First, I investigate whether the baseline characteristics are balanced, that is, those variables determined prior to elections. The Appendix Table A1.6 displays the mean value of each characteristic, difference of means, p-value on the difference, RD estimate, and p-value on RD estimate. I find no statistical evidence of a discontinuous effect at the female win-loss threshold for the baseline covariates. Then I formally test for evidence of sorting around the threshold by implementing the McCrary (2008) density test. The estimate from the test is small and statistically insignificant, showing no evidence of sorting.

1.5.2 Results of Police Responsiveness

Given that law enforcement is one of the most male-dominated public service fields, gender stereotypes that men are linked with leadership activities and women with domestic activities could be very strong (Akerlof and Kranton, 2000). When a female mayor takes office, police might become more tolerant for violence against women as a hostile act.[15] If so, it is hard to construct a scenario in which female victories decrease violence against women. Therefore, it is important to investigate whether female victories lead to deteriorations in local police quality.

Using a sample comprising violent crimes, I estimate the RD model where the dependent variable is an indicator of whether police arrested the offender. Panels A and B of Table 12 display corresponding RD estimates for violence against women and violence against men, respectively. The estimated effect of female mayors on police arrest action is positive and statistically significant without any covariates (see Columns 1 of Panel A). Including a wide variety of covariates at the local and individual level does not alter the results (see Columns 2 to 4). The preferred specification is displayed in Column 4 with full set of control variables. The magnitude of the estimate is approximately 0.14, suggesting that electing female mayors in-

[15]Evidence shows that higher hostile sexism may be linked to more negative evaluations of a female candidate and to lower recommendations that she should be employed (Sakalli-Ugurlu and Beydogan, 2002; Masser and Abrams, 2004).

creases the likelihood of police arresting perpetrators of violence against women by 14 percent. Panel A of Table A1.7 shows robustness across different bandwidths.

It is important to emphasize here that this finding confirms that female victories do not precipitate a backlash with respect to women's safety. In contrast, women occupying leadership positions lead to greater police responsiveness to violence against women. Such a significant effect could also be interpreted as implying that mayors do have direct power over local law enforcement. This is consistent with the argument of Jones (2008) that the development of law enforcement remains tight associations with local electoral processes in the United States. Taken together, these results shed light on the mechanism through which female victories decrease violence against women. The logic I put forth parallels the deterrence theory that an increase in an offender's chances of being caught decreases crime.

Panel B of Table 12 reports the RD estimates for police arrest activity of violence against men. The estimates in Columns 1 through 4 are small and statistically insignificant, suggesting that female victories do not influence police responsiveness to violence against men. Comparing differences in results across genders makes the women-specific channel even more convincing. Notably, it is direct evidence that the improved law enforcement services for women do not come with the sacrifice of such services for men. For comparison purposes, I present graphical illustrations of the discontinued and continuity in arresting for violence against women and men, respectively in Panels A and B of Figure 4. The results from visual inspection confirm the results described above.

1.5.3 Results of Crime Reporting

Having offered detailed evidence on the police responsiveness channel through which electing a female mayor could deter potential offenders from committing crimes against women, now I focus on the reporting decisions of victims, which not only reflect police performance but also influence expected cost of illegal activity.

To uncover the relationship between female mayors and reporting of crimes against women, I use an indicator set to one if the crime was reported to the police as the dependent variable. Panel A of Table 13 displays the corresponding results. I estimate the RD model without control variables in Column 1. Columns 2 through 4 report estimates for examining the sensitivity of the results to additional covariates such as female officer share, total police strength, local economic conditions, and victim characteristics.

I find that the effect of female victories is positive and statistically significant, which is very insensitive to the inclusion of control variables. The preferred specification is displayed in Column 4 with full set of control variables. The magnitude of the coefficient suggests that electing a female mayor leads to an increase in the likelihood of reporting by female victims by approximately 33 percentage points. Panel B of Table A1.7 provides evidence of the robustness of this finding to varying the bandwidth choice.

Importantly, this sizable increase in crime reporting by female victims justifies the large effect of female mayors on homicides against women. It could be thought of as the elasticity for violence against women with respect to crime reporting, where incurring police involvement functions as a key trigger of a response by the criminal justice system and in turn deters potential crime. Moreover, this finding provides additional evidence of an increase in police efforts to protect women after having a female mayor.

There might be spillover effects, as the reporting behavior of male victims could be influenced after they learn that more female victims come forward to report the crime. In contrast to the significant effects of female mayors on reporting decisions of female victims, I do not detect strong effects on reporting decisions of male victims (see Panel B of Table 13). These results strength the previous results for homicides and police responsiveness, shedding light on the channel where women ascending to the highest rung of the governance ladder increase the targeting of benefits to women. As long as female victories do not damage men's perception of credibility of criminal justice system, the net impact of electing female mayors is positive. My findings do indicate that crowding out is not the scenario.

The results, presented in Table 13, are corroborated by visual inspection, as there is a visible discontinuity for reporting of violence against women around the female win-loss threshold (see Panel A of Figure 5). Also, the graphic analysis shows that there is no discontinuity for violence against men, confirming the RD results again (see Panel B of [figure4]Figure 4). Taken together, these results shed light on the crime reporting mechanism, that is, having sympathetic women leaders increases reporting of violence against women, hence triggering behavioral responses by potential offenders. The lack of strong relationship between female mayors and reporting of violence against men further corroborates the gender-specific effect.

In summary, the evidence and falsification checks presented in this section suggest that elections of female mayors increase police responsiveness to violence against women and enhance women's incentives to obtain justice. Importantly, they are consistent with the police responsiveness and crime reporting accounts in which female leadership decreases violence against women.

1.5.4 Placebo Thresholds

As presented in Section 1.4.4, I conduct the same excises to examine the placebo thresholds. Following Imbens and Lemieux (2008), I test jumps at the median of the two subsamples on either side of the cutoff value. Furthermore, I split the sample again at the new median thresholds and set another threshold at the median of the new samples. The corresponding results for arresting and reporting outcomes are presented in Table 14. All the coefficient estimates are insignificant, again suggesting that my results are not due to noisy jumps in the data.

1.6 Discussion

The key argument of my analysis is that potential perpetrators respond to the threat of punishment. In particular, potential offenders of violence against women react when an associated risk of apprehension and subsequent punishment increases after a female mayor takes office. I find that criminals commit less violence against women after a female holds the mayor's office, while I do not find significant changes in violence against men. Importantly, I examine the mechanisms through which female victories affect violence against women. The evidence suggests that female victories induce greater police responsiveness and more reporting that result in the behavioral reduction in crime. This result is important per se, and also to complement the prior literature on women helping women particularly in the criminal justice system.

Some studies point out that violence against women is driven by emotion rather than calculation. For instance, Card and Dahl (2011) show the important role for emotional cues in precipitating violence. Also, Angelucci (2007) demonstrates the association between alcohol consumption and violence.[16] I consider that these studies augment Becker's neoclassical model of offending. Nevertheless, I believe that studying Becker's prediction is as important as it ever was, as underreporting of crime due to social stigma casts doubts on the deterrence value. This study offers evidence that violence against women is sensitive to more crime reporting and greater police responsiveness after female victories, complementing deterrence theory. Nevertheless, future research is needed to design evidence-based crime-reduction interventions with a focus on how emotional cues interact with the threat of punishment.

It is important to reiterate that some recent studies examine the impact of the introduction of Women's Police Stations (WPS) that specialize in serving women on violence against women in Brazil (Perova and Reynolds, 2017), Peru (Kavanaugh et al., 2018), and India (Amaral et al., 2021).[17] However, unlikely those studies, which primarily center on direct interactions between female police officers and female victims in a female-friendly space, this study focuses on the impact of female mayors on the interactions between police, victims, and potential offenders. Nevertheless, in terms of the magnitude, my estimated elasticities are very close to

[16] Also, Chalfin et al. (2021) shows the effect of alcohol consumption on violence during the COVID-19 pandemic in the US.

[17] In Peru, they are named as Women's Justice Centers.

those reported by them. My derived elasticities for homicides against women and reporting of violence against women with respect to female mayors are -1.79 and 0.33, respectively. Correspondingly, for the former the elasticity with respect to WPS is -2.11 (Perova and Reynolds, 2017), while for the latter the elasticities with respect to WPS are 0.29 (Amaral et al., 2021) and 0.40 (Kavanaugh et al., 2018). This means that having a female mayor may have equivalent effects to the opening of WPS.

Police governance and accountability are particularly central to policy and academic debates about police reform. Yet decentralization that aims to strengthen local electoral accountability has rarely been studied as a key determinant of crime in the economics literature. My results show that the identity of local leaders plays an important role in crime. Especially, I find direct evidence that the presence of female mayors leads to greater police responsiveness, suggesting that local governments exert close control over policing in the United States. On the contrary, I find that there is no significant effect of having female governors on crime. These results show the importance of governance arrangements for local law enforcement and crime.

Thus, when it comes to police reform, the discussion should not focus only on individual officer behavior and ignore the full picture of governance structures. I believe that the core content of these findings has implications beyond my setting and context, since some countries with relatively centralized police system have switched to more decentralized ones. For instance, because of the perceived lack of accountability of police authorities to the communities they served, directly elected Police and Crime Commissioners (PCCs) were introduced to England and Wales in 2012 (Sampson, 2012; College of Policing, 2018). Nevertheless, future studies are required to examine the relative effectiveness of different types of police decentralization systems and draw clear inferences regarding police reforms across the world.

1.7 Conclusion

This study uses the lens of gender to explore the role of the mayor's preferences in deterring crime. I offer the first comprehensive analysis of the effect of female mayors on violence against women by exploiting the female win-loss threshold. Importantly, I examine the mechanisms related to police responsiveness and crime reporting by which the identity of mayors influences violence against women.

The following results are documented. First, regression discontinuity estimates show that female victories in mayoral elections decrease violence against women. In contrast, evidence shows that violence against men remains unaffected, ruling out the possibility of generalized effects that influence crime. Second, the crime-reduction effect remains significant throughout the female mayor's term, whereas it fades away after a male succeeds a female as mayor. Third, regardless of victim sex, the impact of female governors on violence is small and statistically insignificant. This result is meaningful and implies the important role of governance arrangement in crime. Fourth, I find that female victories increase police responsiveness to violence against women and reporting by female victims that serve as mechanisms underlying the relationship between female victories and violence against women.

Overall, these results paint a consistent picture of female mayors affecting the behavior of the three categories of interacting agents relevant to a crime situation: police, victims, and criminals. Despite the importance of this study, further research is needed to offer more insight into the effective ways to reach the goal of sustainable representation of women in political leadership.

Table 1: Summary Statistics

	(1)	(2)	(3)
	N	Mean	S.D.
Panel A: Homicides per 100,000 Population			
Female Victims Only	1,695	2.59	1.80
Male Victims Only	1,695	11.93	10.63
Female and Male Victims	1,695	0.17	0.27
	N	Fraction	
Panel B: Female Participation			
Total Elections	4,124		
Mixed-Gender Elections	812	0.20	
	N	Mean	S.D.
Panel C: Vote Share in Mixed-Gender Elections			
Female Vote Share	812	0.48	0.16
Male Vote Share	812	0.47	0.17
Female Margin of Victory	812	0.01	0.31

Notes: This table displays summary statistics of homicides and mayoral elections in the sample used for the main analysis. Panel A displays observations, means, standard deviations of homicides by victim sex after mixed-gender elections. Panel B presents the numbers of total elections and mixed-gender elections, and the share of mixed-gender elections. Panel C displays observations, means, standard deviations of vote share by gender and female margin of victory.

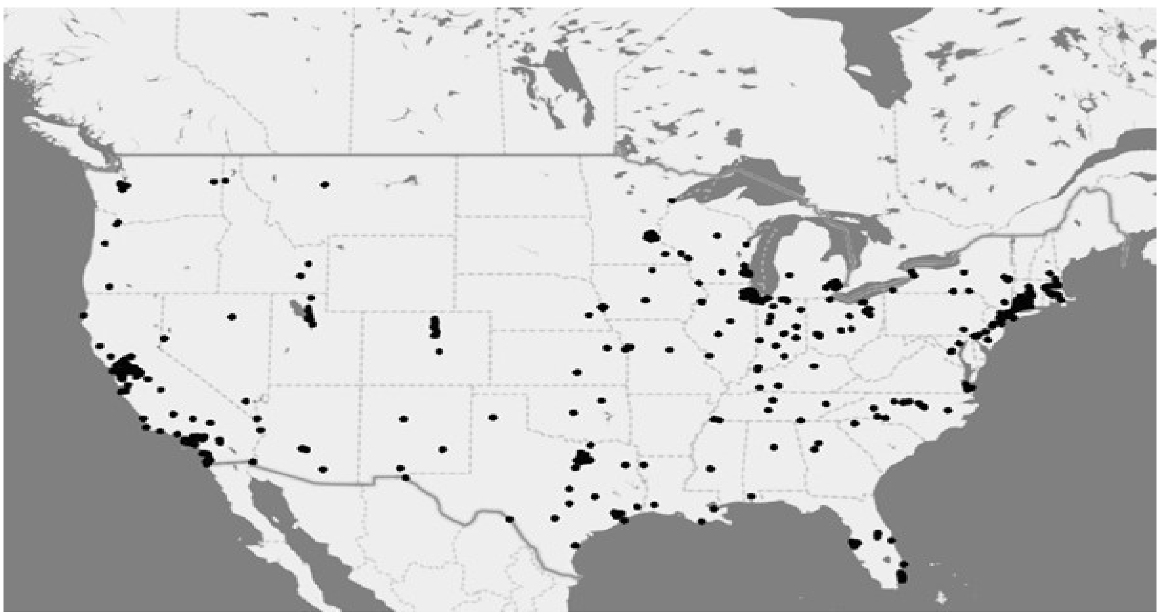


Fig. 1: Mixed-Gender Election Cities

Table 2: Baseline Characteristics

	(1)	(2)	(3)	(4)	(5)	(6)
	7 percent vote spread					
	Female won	Female lost	Difference of means	p-value on difference	RD estimate	p-value on RD estimate
Population density	1,393.07	1,068.69	324.38	0.12	163.53	0.76
Income per capita (in $1,000)	15.67	14.87	0.80	0.48	-0.43	0.79
Unemployed (%)	7.22	7.05	0.17	0.74	0.69	0.52
Land area (km^2)	169.14	168.68	0.46	0.99	137.13	0.44
Water area (km^2)	17.20	6.67	10.53	0.29	-9.19	0.54
Elevation (m)	211.33	255.07	-43.74	0.41	53.54	0.68
Female incumbent	0.34	0.22	0.12	0.12	0.22	0.16
Police officers per 100,000 persons	193.79	177.05	16.74	0.31	36.75	0.44
Female officer share	0.07	0.06	0.01	0.08	-0.00	0.79
Police civilians per 100,000 persons	64.69	54.51	10.18	0.50	26.73	0.60
Female civilian share	0.71	0.70	0.01	0.75	0.07	0.29
Public employees per 1,000 persons	17.85	21.39	-3.54	0.15	3.42	0.56
March payroll per public employee ($)	2,488.80	2,355.89	132.91	0.41	-336.03	0.35
Observations	76	63	139		139	

Notes: This table examines whether pre-characteristics are balanced across female win-loss threshold. Columns 1 and 2 show the unconditional means for cities where the females barely won and where they barely lost. Column 3 shows the difference of means across columns 1 and 2, and column 4 shows the p-value for the difference of means. Columns 5 reports the estimate on female win from a RD specification where the respective characteristic is used as the dependent variable. Column 6 is the p-value for a RD estimate. An optimal bandwidth of 0.07 around the threshold has been used.

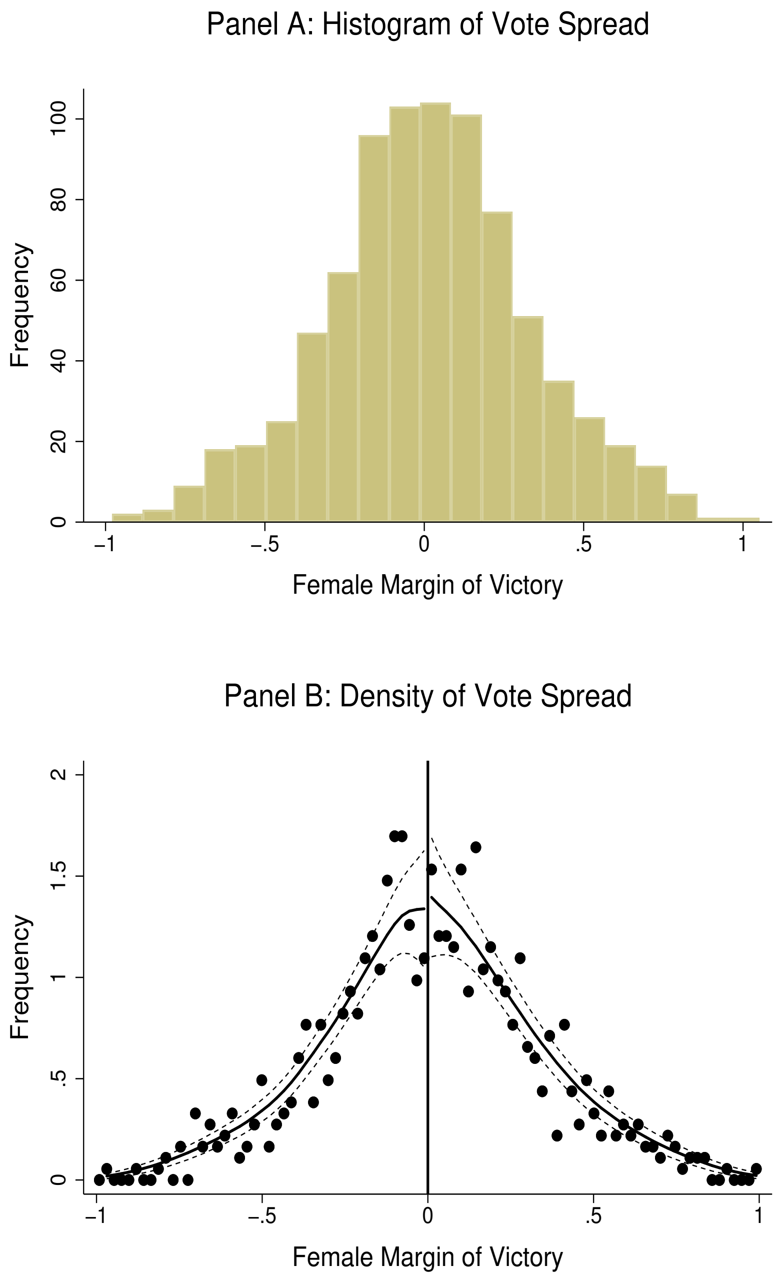


Fig. 2: Homicides and Female Margin of Victory

Notes: The figure shows the distribution of vote spread. Panel A shows the histogram of vote spread. Panel B shows the density of vote spread. The solid line plots predicted values from a local linear regression of frequency on vote spread, with separate vote spread trends estimated on either side of the female win-loss threshold, following McCrary (2008). The dashed lines show 95% confidence intervals. The McCrary test statistic is 0.05 (standard error 0.18).

Table 3: RD Estimates of the Effect of Female Mayors on Homicides

	(1)	(2)	(3)	(4)
	Homicide Rate			
Panel A: Female Victims Only				
Female Win	-2.234**	-2.016**	-1.998**	-1.791**
	(0.941)	(0.838)	(0.843)	(0.779)
City- and Year-Fixed Effects	YES	YES	YES	YES
Female Officer Share	NO	YES	YES	YES
Police Forces	NO	NO	YES	YES
Other Controls	NO	NO	NO	YES
Observations	345	345	345	345
Panel B: Male Victims				
Female Win	-6.552	-7.141	-6.878	-4.415
	(5.159)	(5.138)	(4.588)	(4.014)
City- and Year-Fixed Effects	YES	YES	YES	YES
Female Officer Share	NO	YES	YES	YES
Police Forces	NO	NO	YES	YES
Other Controls	NO	NO	NO	YES
Observations	345	345	345	345

Notes: This table presents RD estimates of the effect of being above the female win-loss threshold on homicides per 100,000 persons. Panels A and B display estimates for homicides involving female victims only and male victims, respectively. Other controls include unemployment, income per capita, and population. Standard errors, clustered at the city level, are reported in parenthesis. Significance at the 10%, 5%, and 1% levels is indicated by *, **, and ***, respectively. An optimal bandwidth of 0.07 around the female win-loss threshold has been used.

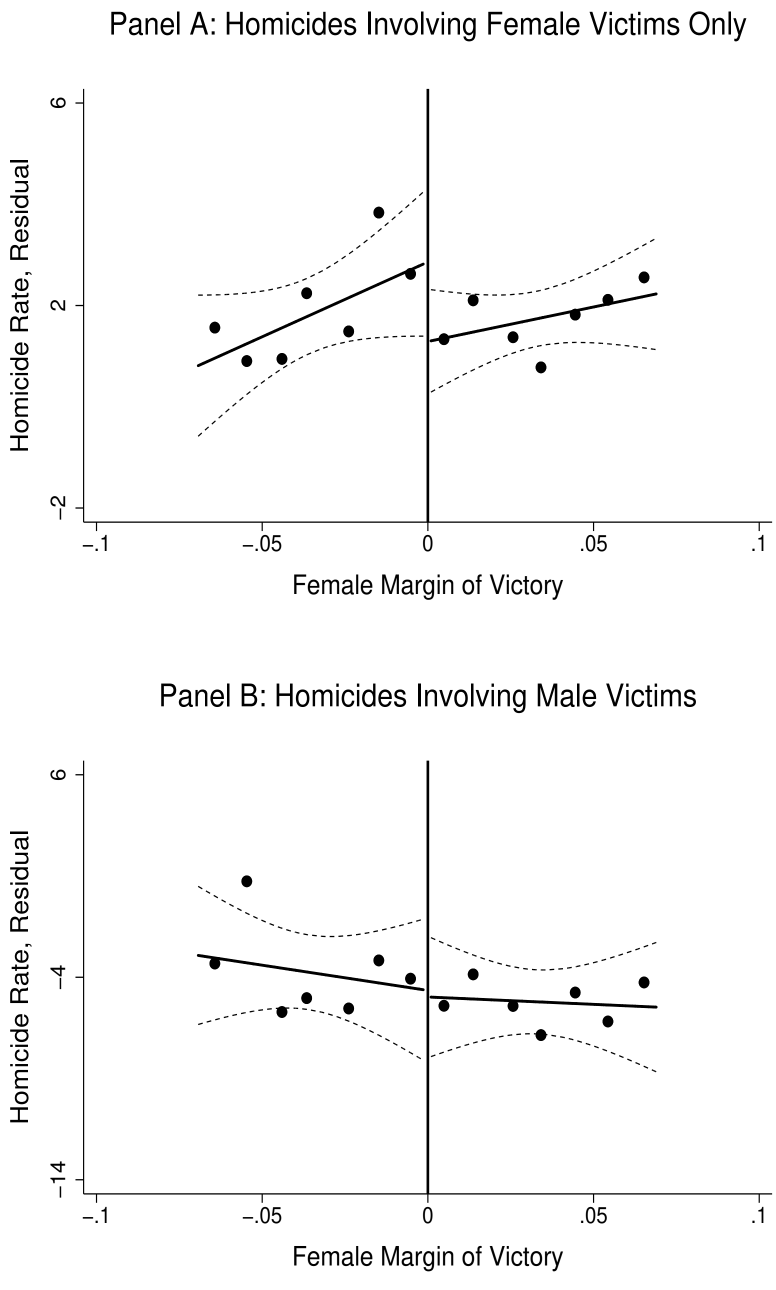


Fig. 3: Homicides and Female Margin of Victory

Notes: This figure plots the residualised values after controlling for all covariates in the main specification of homicides per 100,000 persons against the margin of victory, with a negative margin indicating a female candidate loss. Panels A and B display estimates for homicides involving female victims only and male victims, respectively. Each point represents the average value of the outcome in the margin of victory bin. The solid line plots predicted values, with separate vote spread trends estimated on either side of the female win-loss threshold, while the dashed lines identifies the 95 percent confidence level.

Table 4: RD Estimates of the Effect of Female Mayors on Pre-Existing Homicides

	(1)	(2)	(3)	(4)
	Homicide Rate			
	Female Victims Only		Male Victims	
Female Win	-0.220	0.206	-3.150	-2.487
	(1.045)	(0.799)	(2.800)	(1.851)
City- and Year-Fixed Effects	YES	YES	YES	YES
Controls	NO	YES	NO	YES
Observations	323	323	323	323

Notes: This table presents RD estimates of the effect of female mayors on pre-existing homicides by gender of victims. Columns 1 and 2 report estimates for homicides involving female victims only, while Columns 3 and 4 display estimates for homicides involving male victims. Other controls include unemployment, income per capita, and population. Standard errors, clustered at the city level, are reported in parenthesis. Significance at the 10%, 5%, and 1% levels is indicated by *, **, and ***, respectively..

Table 5: Placebo Tests Using Alternative Thresholds

	(1)	(2)	(3)
Thresholds	Coefficients	Standard Errors	Observations
Panel A			
-0.18	0.07	0.42	868
0.19	-0.14	0.27	847
Panel B			
-0.34	-0.14	0.43	444
-0.10	1.22	0.90	424
0.10	1.03	0.87	452
0.35	-0.96	0.80	395

Notes: Notes: The table shows RD estimates of the effect of placebo thresholds on homicides against women for the election subsample. The following female margin of victory ranges are used: [-0.97, 0), (0, 0.98], [-0.97, -0.18), [-0.18, 0), (0, 0.19], and (0.19, 0.98]. In each range a placebo RD threshold is created at the median.

Table 6: RD Estimates of the Effect of Female Mayors on Homicides by Year

	(1)	(2)	(3)	(4)
	Homicide Rate			
	Female Victims Only		Male Victims	
Female Win × First Year	-2.666***	-2.296***	-6.586	-3.779
	(0.974)	(0.815)	(5.767)	(4.834)
Female Win × Second Year	-2.273**	-1.934**	-6.723	-3.737
	(0.947)	(0.752)	(5.839)	(4.592)
Female Win × Third Year	-2.211**	-1.940**	-7.290	-4.648
	(1.079)	(0.771)	(5.404)	(4.122)
Female Win × Fourth Year	-2.235**	-1.908**	-5.622	-2.890
	(1.061)	(0.872)	(5.120)	(4.091)
City- and Year-Fixed Effects	YES	YES	YES	YES
Controls	NO	YES	NO	YES
Observations	345	345	345	345

Notes: This table presents RD estimates of the effect of female mayors by year of the mayor's term on homicide rate. Columns 1 and 2 report estimates for homicides involving female victims only, while Columns 3 and 4 display estimates for homicides involving male victims. Other controls include unemployment, income per capita, and population. Standard errors, clustered at the city level, are reported in parenthesis. Significance at the 10%, 5%, and 1% levels is indicated by *, **, and ***, respectively.

Table 7: RD Estimates of the Effect of Female Mayors on Homicides in the Subsequent Mayor Term

	(1)	(2)	(3)	(4)
	Homicide Rate			
	Female Victims Only		Male Victims	
Female Win	-0.939		1.884	
	(1.634)		(4.786)	
Female Win × First Year		-0.945		1.004
		(1.686)		(5.018)
Female Win × Second Year		-1.316		1.761
		(1.670)		(5.189)
Female Win × Third Year		-0.867		1.924
		(1.675)		(5.281)
Female Win × Fourth Year		-0.303		1.288
		(1.653)		(5.422)
Observations	339	339	339	339

Notes: This table shows RD estimates of the effect of female mayors on homicides in the subsequent mayor term. The first two columns report estimates for homicides involving female victims only, while the last two columns display estimates for homicides involving male victims. Other controls include unemployment, income per capita, and population. Standard errors, clustered at the city level, are reported in parenthesis. Significance at the 10%, 5%, and 1% levels is indicated by *, **, and ***, respectively.

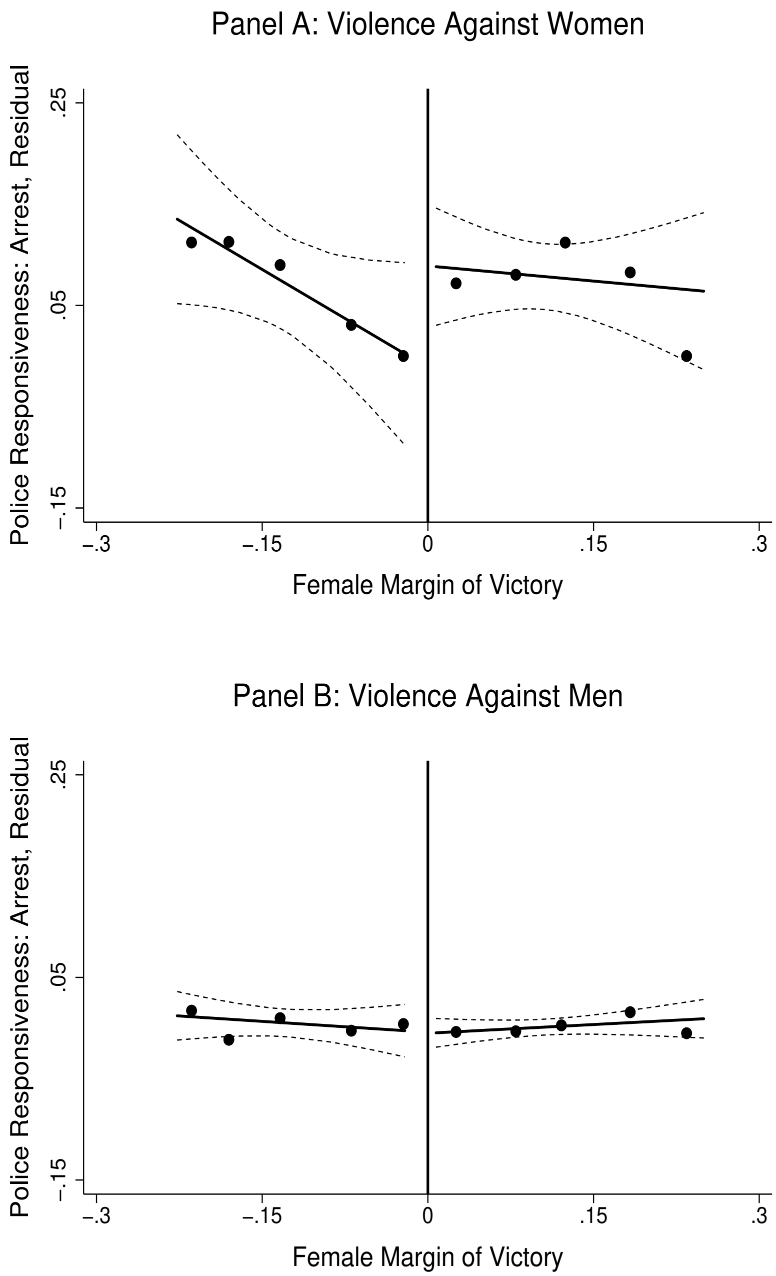
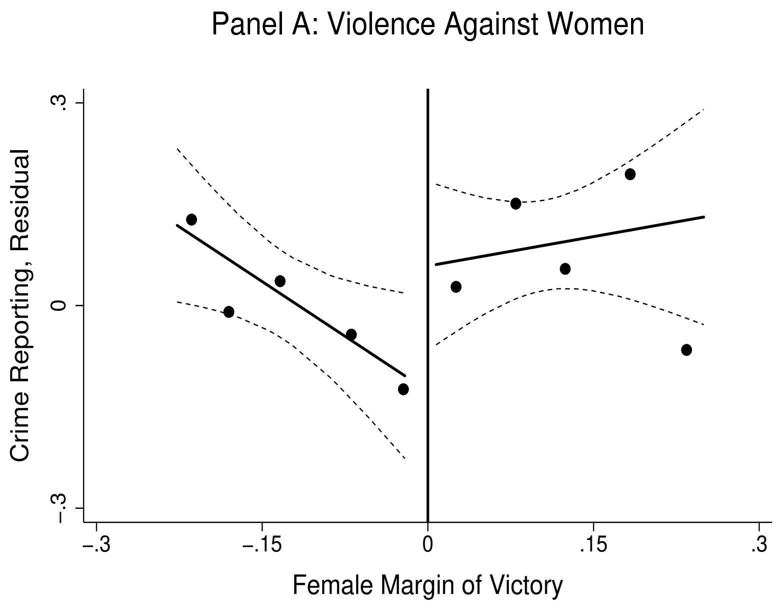


Fig. 4: Police Responsiveness and Female Margin of Victory

Notes: This figure plots the residualised values after controlling for all covariates in the main specification of police arresting against the margin of victory, with a negative margin indicating a female candidate loss. Panels A and B display estimates for violence against women and violence against men, respectively. Each point represents the average value of the outcome in the margin of victory bin. The solid line plots predicted values, with separate vote spread trends estimated on either side of the female win-loss threshold, while the dashed lines identifies the 95 percent confidence level.



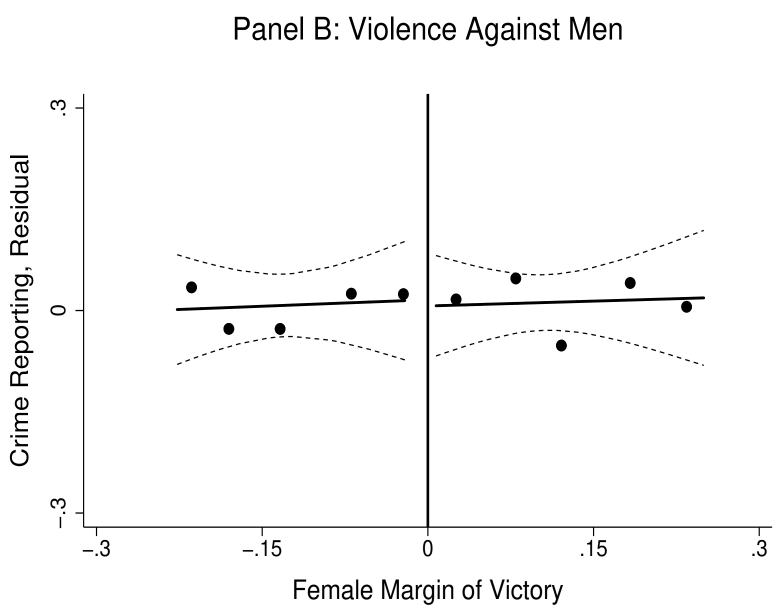


Fig. 5: Crime Reporting and Female Margin of Victory

Notes: This figure plots the residualised values after controlling for all covariates in the main specification of crime reporting against the margin of victory, with a negative margin indicating a female candidate loss. Panels A and B display estimates for violence against women and violence against men, respectively. Each point represents the average value of the outcome in the margin of victory bin. The solid line plots predicted values, with separate vote spread trends estimated on either side of the female win-loss threshold, while the dashed lines identifies the 95 percent confidence level.

Table 8: RD Estimates of the Effect of Female Mayors on Probability of at Least One Homicide Occurring

	(1)	(2)	(3)	(4)
	Homicide Probability			
	Female Victims Only		Male Victims	
Female Win	-0.513**	-0.483**	0.085	0.069
	(0.214)	(0.216)	(0.152)	(0.159)
City- and Year-Fixed Effects	YES	YES	YES	YES
Controls	NO	YES	NO	YES
Observations	345	345	345	345

Notes: This table presents RD estimates of the effect of female mayors on probability of at least one homicide occurring. Columns 1 and 2 show estimates for homicides involving female victims only, while Columns 3 and 4 display estimates for homicides involving male victims. Other controls include unemployment, income per capita, and population. Standard errors, clustered at the city level, are reported in parenthesis. Significance at the 10%, 5%, and 1% levels is indicated by *, **, and ***, respectively.

Table 9: RD Estimates of the Effect of Female Governors on Homicides

	(1)	(2)	(3)	(4)
	Homicide Rate			
	Female Victims Only		Male Victims	
Female Governor Win	-0.115	0.170	-2.087	-2.669
	(0.348)	(0.450)	(1.589)	(1.926)
State- and Year-Fixed Effects	YES	YES	YES	YES
Controls	NO	YES	NO	YES
Observations	139	139	139	139

Notes: This table presents RD estimates of the effect of female governors on homicides by gender of victims. Columns 1 and 2 show estimates for homicides involving female victims only, while Columns 3 and 4 display estimates for homicides involving male victims. Standard errors, clustered at the state level, are reported in parenthesis. Significance at the 10%, 5%, and 1% levels is indicated by *, **, and ***, respectively. An optimal bandwidth around the threshold has been used.

Table 10: RD Estimates of the Effect of Female Mayors on Domestic Homicides

	(1)	(2)	(3)	(4)
	Domestic Homicides			
Female Win	-1.174**	-1.099**	-1.100**	-1.128**
	(0.492)	(0.498)	(0.501)	(0.484)
City- and Year-Fixed Effects	YES	YES	YES	YES
Female Officer Share	NO	YES	YES	YES
Police Forces	NO	NO	YES	YES
Other Controls	NO	NO	NO	YES
Observations	345	345	345	345

Notes: This table presents RD estimates of the effect of being above the female win-loss threshold on domestic homicides against women per 100,000 persons. Other controls include unemployment, income per capita, and population. Standard errors, clustered at the city level, are reported in parenthesis. Significance at the 10%, 5%, and 1% levels is indicated by *, **, and ***, respectively. An optimal bandwidth around the female win-loss threshold has been used.

Table 11: RD Estimates of the Effect of Democratic Mayors on Homicides against Women

	(1)	(2)	(3)	(4)
	Homicide Rate			
	Homicides against Women		Domestic Homicides	
Democratic Win	-0.121	0.039	-0.255	-0.171
	(0.580)	(0.391)	(0.170)	(0.153)
City- and Year-Fixed Effects	YES	YES	YES	YES
Controls	NO	YES	NO	YES
Observations	621	617	621	617

Notes: This table presents RD estimates of the effect of Democratic mayors on homicides against women. Columns 1 and 2 show estimates for homicides against women, while Columns 3 and 4 display estimates for domestic homicides. Other controls include population, female officer share, and total police forces. Standard errors, clustered at the city level, are reported in parenthesis. Significance at the 10%, 5%, and 1% levels is indicated by *, **, and ***, respectively.

Table 12: RD Estimates of the Effect of Female Mayors on Police Responsiveness

	(1)	(2)	(3)	(4)
	Did the Police Make Arrest? (Yes = 1)			
Panel A: Violence Against Women				
Female Win	0.169**	0.180**	0.188**	0.139*
	(0.066)	(0.066)	(0.068)	(0.069)
Area- and Year-Fixed Effects	YES	YES	YES	YES
Female Officer Share	NO	YES	YES	YES
Police Forces	NO	NO	YES	YES
Other Controls	NO	NO	NO	YES
Observations	1161	1161	1161	1161
Panel B: Violence Against Men				
Female Win	-0.020	-0.018	-0.010	-0.040
	(0.047)	(0.049)	(0.044)	(0.056)
Area- and Year-Fixed Effects	YES	YES	YES	YES
Female Officer Share	NO	YES	YES	YES
Police Forces	NO	NO	YES	YES
Other Controls	NO	NO	NO	YES
Observations	1355	1355	1355	1355

Notes: This table presents RD estimates of the effect of being above the female win-loss threshold on the probability of police apprehending the offender. Panels A and B display estimates for violence against women and violence against men, respectively. Local controls include female officer share, total police forces, unemployment, and income per capita. Victim characteristics include race, Hispanic origin, marital status, and age. Standard errors, clustered at the area level, are reported in parenthesis. Significance at the 10%, 5%, and 1% levels is indicated by *, **, and ***, respectively. An optimal bandwidth of 0.25 around the female win-loss threshold has been used.

Table 13: RD Estimates of the Effect of Female Mayors on Crime Reporting

	(1)	(2)	(3)	(4)
	Was the Crime Reported to Police? (Yes = 1)			
Panel A: Violence Against Women				
Female Win	0.527***	0.531***	0.542***	0.329**
	(0.077)	(0.078)	(0.090)	(0.148)
Area- and Year-Fixed Effects	YES	YES	YES	YES
Female Officer Share	NO	YES	YES	YES
Police Forces	NO	NO	YES	YES
Other Controls	NO	NO	NO	YES
Observations	1161	1161	1161	1161
Panel B: Violence Against Men				
Female Win	-0.083	-0.094	-0.046	0.040
	(0.135)	(0.137)	(0.153)	(0.164)
Area- and Year-Fixed Effects	YES	YES	YES	YES
Female Officer Share	NO	YES	YES	YES
Police Forces	NO	NO	YES	YES
Other Controls	NO	NO	NO	YES
Observations	1355	1355	1355	1355

Notes: This table presents RD estimates of the effect of being above the female win-loss threshold on the probability that the victim reports the crime to the police. Panels A and B display estimates for violence against women and male violence against men, respectively.Local controls include female officer share, total police forces, unemployment, and income per capita. Victim characteristics include race, Hispanic origin, marital status, and age. Standard errors, clustered at the area level, are reported in parenthesis. Significance at the 10%, 5%, and 1% levels is indicated by *, **, and ***, respectively.

Table 14: Placebo Tests Using Alternative Thresholds

	(1)	(2)	(3)
Thresholds	Coefficients	Standard Errors	Observations
Panel A: Arresting			
-0.27	0.10	0.07	938
0.24	0.03	0.07	1,367
-0.37	0.04	0.03	373
-0.15	-0.06	0.04	565
0.12	0.03	0.06	675
0.55	-0.00	0.03	692
Panel B: Reporting			
-0.27	0.12	0.12	938
0.24	0.02	0.15	1,367
-0.37	-0.29	0.20	373
-0.15	0.00	0.06	565
0.12	0.03	0.08	675
0.55	-0.07	0.08	692

Notes: Notes: The table shows RD estimates of the effect of placebo thresholds on arresting and reporting outcomes for the election subsample in Panel A and B, respectively. The following female margin of victory ranges are used: [-0.79, 0), (0, 0.85], [-0.79, -0.27), [-0.27, 0), (0, 0.24], and (0.24, 0.85]. In each range a placebo RD threshold is created at the median.

Table A1.1: Summary Statistics

	(1)	(2)	(3)
	N	Mean	S.D.
Panel A: Violent Crime Against Women (Dummy)			
Police Make Arrest	2,334	0.05	0.21
Crime Reported to the Police	2,334	0.42	0.49
	N	Mean	S.D.
Panel B: Violent Crime Against Men (Dummy)			
Police Make Arrest	3,003	0.06	0.24
Crime Reported to the Police	3,003	0.40	0.49
	N	Fraction	
Panel C: Female Participation			
Total Elections	271		
Mixed-Gender Elections	86	0.32	
	N	Mean	S.D.
Panel D: Vote Share in Mixed-Gender Elections			
Female Vote Share	86	0.50	0.18
Male Vote Share	86	0.44	0.19
Female Margin of Victory	812	0.06	0.36

Notes: This table displays summary statistics of two main outcome variables and mayoral elections matched to crime survey data from 1979 to 2004. Panel A and B display observations, means, standard deviations of police responsiveness and crime reporting for violence against women and men, respectively. Panel C presents the numbers of total elections and mixed-gender elections, and the share of mixed-gender elections. Panel D displays observations, means, standard deviations of vote share by gender, and female margin of victory.

Table A1.2: Robustness of Estimates to Varying Bandwidths

	(1)	(2)	(3)	(4)	(5)
	0.10	0.15	0.30	0.45	0.60
Female Win	-1.037***	-0.928**	-0.751**	-0.447*	-0.507**
	(0.349)	(0.379)	(0.349)	(0.241)	(0.252)
City- and Year-Fixed Effects	YES	YES	YES	YES	YES
Female Officer Share	YES	YES	YES	YES	YES
Police Forces	YES	YES	YES	YES	YES
Other Controls	YES	YES	YES	YES	YES
Observations	527	757	1,172	1,437	1,579

Notes: This table reports RD estimates of the effect of female mayors on homicides involving female victims only across various bandwidths. See the notes to Table 3 for additional details.

Table A1.3: RD Estimates of the Effect of Female Mayors on Homicides Involving Female Victims Only: A Quadratic Polynomial

	(1)	(2)	(3)	(4)
	Homicide Rate			
Female Win	-1.990*	-1.885**	-1.869**	-1.663*
	(1.012)	(0.935)	(0.934)	(0.887)
City- and Year-Fixed Effects	YES	YES	YES	YES
Female Officer Share	NO	YES	YES	YES
Police Forces	NO	NO	YES	YES
Other Controls	NO	NO	NO	YES
Observations	345	345	345	345

Notes: This table presents RD estimates of the effect of female mayors on homicides involving female victims only with a quadratic polynomial. Other controls include unemployment, income per capita, and population. Standard errors, clustered at the city level, are reported in parenthesis. Significance at the 10%, 5%, and 1% levels is indicated by *, **, and ***, respectively.

Table A1.4: RD Estimates of the Effect of Female Mayors on Homicides Involving Female and Male Victims

	(1)	(2)	(3)	(4)
	Homicide Rate			
Female Win	0.250	0.281	0.278	0.186
	(0.347)	(0.356)	(0.361)	(0.295)
City- and Year-Fixed Effects	YES	YES	YES	YES
Female Officer Share	NO	YES	YES	YES
Police Forces	NO	NO	YES	YES
Other Controls	NO	NO	NO	YES
Observations	345	345	345	345

Notes: This table presents RD estimates of the effect of female mayors on homicides involving female and male victims. Other controls include unemployment, income per capita, and population. Standard errors, clustered at the city level, are reported in parenthesis. Significance at the 10%, 5%, and 1% levels is indicated by *, **, and ***, respectively.

Table A1.5: RD Estimates of the Effect of Female Mayors on Homicides (Excl. Election Years)

	(1)	(2)	(3)	(4)
	Homicide Rate			
	Female Victims Only		Male Victims	
Female Win	-3.137*	-3.311**	-4.490	-3.901
	(1.685)	(1.563)	(5.228)	(5.605)
City- and Year-Fixed Effects	YES	YES	YES	YES
Controls	NO	YES	NO	YES
Observations	228	228	228	228

Notes: This table presents RD estimates of the effect of female mayors on homicides excluding the sample in election years. Columns 1 and 2 report estimates for homicides involving female victims only, while Columns 3 and 4 display estimates for homicides involving male victims. Other controls include unemployment, income per capita, and population. Standard errors, clustered at the city level, are reported in parenthesis. Significance at the 10%, 5%, and 1% levels is indicated by *, **, and ***, respectively.

Table A1.6: Baseline Characteristics

	(1)	(2)	(3)	(4)	(5)	(6)
	25 percent vote spread					
	Female won	Female lost	Difference of means	*p*-value on difference	RD estimate	*p*-value on RD estimate
Population density	1,501.22	2,340.52	-839.30	0.10	-3,203.66	0.28
Income per capita (in $1,000)	16.43	15.25	1.18	0.48	-2.60	0.54
Unemployed (%)	6.88	8.04	-1.16	0.20	-0.08	0.98
Land area (km^2)	540.47	532.67	7.80	0.95	-493.45	0.22
Water area (km^2)	101.23	46.21	55.02	0.27	-98.75	0.63
Elevation (m)	90.53	208.81	-118.28	0.12	177.53	0.28
Female incumbent	0.33	0.18	0.15	0.27	-0.30	0.39
Police officers per 100,000 persons	226.17	268.34	-42.17	0.31	-203.50	0.22
Female officer share	0.11	0.11	0.00	0.92	-0.07	0.14
Police civilians per 100,000 persons	76.94	73.98	2.96	0.82	-58.07	0.17
Female civilian share	0.73	0.70	0.03	0.35	0.03	0.73
Public employees per 1,000 persons	17.14	20.20	-3.06	0.47	-25.11	0.24
March payroll per public employee ($)	2,585.75	2,743.67	-157.92	0.52	21.92	0.95
Observations	27	17	44		44	

Notes: This table examines whether pre-characteristics are balanced across female win-loss threshold. Columns 1 and 2 show the unconditional means for cities where the females barely won and where they barely lost. Column 3 shows the difference of means across columns 1 and 2, and column 4 shows the *p*-value for the difference of means. Columns 5 reports the estimate on female win from a RD specification where the respective characteristic is used as the dependent variable. Column 6 is the *p*-value for a RD estimate. An optimal bandwidth of 0.25 around the threshold has been used.

Table A1.7: Robustness of Estimates to Varying Bandwidths

	(1)	(2)	(3)	(4)	(5)	(6)
	0.15	0.20	0.25	0.30	0.35	0.40
Panel A: Arresting						
Female Win	0.302***	0.079*	0.139*	0.155***	0.104***	0.058*
	(0.047)	(0.043)	(0.069)	(0.034)	(0.036)	(0.029)
Area- and Year-Fixed Effects	YES	YES	YES	YES	YES	YES
Female Officer Share	YES	YES	YES	YES	YES	YES
Police Forces	YES	YES	YES	YES	YES	YES
Other Controls	YES	YES	YES	YES	YES	YES
Observations	936	1069	1161	1382	1584	1637
Panel B: Reporting						
Female Win	0.364**	0.320**	0.329**	0.233***	0.232***	0.127*
	(0.165)	(0.148)	(0.148)	(0.066)	(0.079)	(0.070)
Area- and Year-Fixed Effects	YES	YES	YES	YES	YES	YES
Female Officer Share	YES	YES	YES	YES	YES	YES
Police Forces	YES	YES	YES	YES	YES	YES
Other Controls	YES	YES	YES	YES	YES	YES
Observations	936	1069	1161	1382	1584	1637

Notes: This table reports RD estimates of the effect of female mayors on police responsiveness to violence against women and reporting of violence against women across various bandwidths. See the notes to Table 9 for additional details.

Table A1.8: RD Estimates of the Effect of Female Mayors: A Quadratic Polynomial

	(1)	(2)	(3)	(4)
Panel A: Arresting				
Female Win	0.506***	0.503***	0.508***	0.238*
	(0.083)	(0.093)	(0.097)	(0.135)
Area- and Year-Fixed Effects	YES	YES	YES	YES
Female Officer Share	NO	YES	YES	YES
Police Forces	NO	NO	YES	YES
Other Controls	NO	NO	NO	YES
Observations	1161	1161	1161	1161
Panel B: Reporting				
Female Win	0.344***	0.342***	0.343***	0.286***
	(0.045)	(0.046)	(0.045)	(0.057)
Area- and Year-Fixed Effects	YES	YES	YES	YES
Female Officer Share	NO	YES	YES	YES
Police Forces	NO	NO	YES	YES
Other Controls	NO	NO	NO	YES
Observations	1161	1161	1161	1161

Notes: This table presents RD estimates of the effect of being above the female win-loss threshold with a quadratic polynomial. Panel A displays estimates for police responsiveness to violence against women, while Panel B shows estimates for reporting of violence against women. Local controls include female officer share, total police forces, unemployment, and income per capita. Victim characteristics include race, Hispanic origin, marital status, and age. Standard errors, clustered at the area level, are reported in parenthesis. Significance at the 10%, 5%, and 1% levels is indicated by *, **, and ***, respectively.

CHAPTER 2

Female Chief Officers and Crime: Evidence from England and Wales

Abstract: I study the impact of the appointment of female chief officers in policing on female salient crimes: sexual and rape offences. Evidence shows that appointing more female chief officers leads to a statistically significant increase in recorded female salient crimes in England and Wales. Yet, this rise is good news, which is due not to a rise in actual crimes committed but, rather, to an increase in essential inputs into the production of law enforcement services. I also find that women's ascension to top positions in law enforcement decreases violence against women. These findings are further corroborated by the event study analysis for the appointment of the first female chief officers.

2.1 Introduction

Worldwide, 35 percent of women have experienced sexual violence, has been increasingly recognized as a global public health problem (García-Moreno et al., 2013). Voluminous studies show far-reaching consequences of abusing women, especially in regard to poor physical health (Ulrich et al., 2003), poor mental health (Plichta and Falik, 2001; Oram et al., 2017), low birth weight (Aizer, 2011; Neggers et al., 2004), and child malnutrition (Sobkoviak et al., 2012). In line with the 2030 Agenda for Sustainable Development, the European Union and United Nations are embarking on the Spotlight Initiative as the world's largest targeted effort to end all forms of violence against women, where women's empowerment is the central to achieve the goal (Cebula and Salsberg, 2019).

Women's empowerment defined as improving the ability of women to access the constituents of development—in particular health, education, earning opportunities, rights, and political participation (Duflo, 2012), consists mostly of economic and political empowerment. In fact, many scholars have examined the relationships between economically empowering women by increasing employment opportunity and income of women and women's exposure to domestic violence (Bhalotra et al., 2020; Anderberg et al., 2016; Bobonis et al., 2013; Aizer, 2010; Hidrobo and Fernald, 2013). Yet at the same time, the results of some studies also indicate that promoting economic empowerment of women with low bargaining power may put them at more risk of Intimate partner violence, as men seek to counteract their increased bargaining power (Rahman et al., 2011; Friedemann-Sánchez and Lovatón, 2012; Heath, 2014).

Apart from the intra-household bargaining channels, political empowerment of women often plays a role in reducing sexual crimes through more channels. Based on the evidence of the striking difference in political and policy preferences between men and women(Lott and Kenny, 1999; Inglehart and Norris, 2000; Dolan, 2000; Edlund and Pande, 2002), a number of papers have already concentrated on how political reservations for women increase the targeting of benefits to women (Powley, 2007; Clots-Figueras, 2011), going back to the seminal contribution of Chattopadhyay and Duflo (2004).[1] Moreover, women holding political positions may serve as "role models" for women in the population, which echoes the economic theory of mentoring and diversity (Athey et al., 2000).

[1] A separate branch of this literature investigates the impact of mandated political representation in providing disadvantaged minorities influence over policy-making (Pande, 2003; Besley et al., 2012). However, more recent studies have questioned the aforementioned assertions (Bardhan et al., 2010; Rajaraman and Gupta, 2012; Dunning and Nilekani, 2013; Ferreira and Gyourko, 2014).

There is a dense literature in the applications of the theories developed in Chattopadhyay and Duflo (2004) and Athey et al. (2000) in the context of crimes against women. For instance, Iyer et al. (2012) provide empirical evidence that increased representation among members of the village council could influence the functioning of the police and, in turn, make police take women's grievances more seriously. As a result, victims are more encouraged to report sexual crimes. Similarly, Anwar et al. (2019) find that adding females to the jury pool significantly impacted conviction rates on sexual offence cases. However, using data from 583 Indian districts from 2002 to 2007, Sekhri and Storeygard (2014) do not find supportive evidence for mitigating effects of women's political representation in the national parliament on dowry deaths.

A small number of studies focus on the implications of female leadership in politics for sexual crimes.[2] As described above, female leaders are more likely to set priorities and allocate resources that are directly relevant to the needs of their own genders. However, in contrast to most research on the influence of female leadership, Iyer et al. (2012) find no effect of gender of council leaders at the district level on the occurrences of crimes against women. A possible explanation for this insignificant effect could lie in a centralised system of the Indian Police Services where officers are appointed by the federal government (Lambert et al., 2015) and women's presence in leadership positions at the district level have no direct power over police appointments, budgets or strategies.

The existing literature also investigates the direct effect of female representation in law enforcement on sexual crimes. Considering that female officers have a distinctive advantage in interactions with female assault victims (Thielemann and Stewart Jr, 1996; Harrington et al., 2003; Riccucci et al., 2014; Miller and Segal, 2019),[3] the literature show that greater female police representation could increase

[2] Related is also the literature on the race of leaders, such as Bulman (2019).

[3] Many studies have indicated that men display a significantly higher endorsement of rape myths than women (Anderson et al., 1997; Suarez and Gadalla, 2010). Meanwhile, men tend to attribute less blame in sexual assault to perpetrators than do women (Alicke and Yurak, 1995; García-Moreno et al., 2013). However, Wentz and Archbold (2012) and Morabito et al. (2017) do not find support for the above-mentioned results.

the reporting rates of violence against women and decrease the incidence.

Additionally, concerned governments have implemented an innovative policy intervention of women's police stations aimed at eradicating crimes against women where only female officers deal directly with survivors and women's representation in policing rapidly increases. Enhancing women's access to justice, the introduction of women's police stations improves their willingness to report crimes but also lowers incidence of homicide for female victims (Perova and Reynolds, 2017; Kavanaugh et al., 2018; Amaral et al., 2021).

This paper contributes to the existing literature by bringing together two streams of literature: police leadership and gender-based crimes, adding something new to the picture. Because sexual violence is an issues of particular concern for women, for instance, over 83 percent and 95 percent of the victims of sexual offences and rape offences, respectively are female in England and Wales (Flatley, 2018),[4] it allows an opportunity to explore the impact of female leaders in law enforcement on a gendered issue. I investigate thoroughly whether female chief officers influence female salient crimes recorded: sexual offences and rape offences. Also, collecting data on homicides against women, I explore the impact of female chief officers on actual incidence of violence against women.

In doing so, my study diverges from existing literature in two main ways. For one, to the best of my knowledge, this is the first paper to analyse the effect of adding female chief officers in the absence of a dramatic growth in women's representation across ranks of police due to affirmative action quotas or women's police stations. Over the period between 2005 and 2019, the number of female chief officers in England and Wales nearly tripled from 22 to 58, while that of male counterparts declined from 192 to 159.[5] Yet little is known about the potentially important role of female leaders in one of the most male-dominated public service fields.

Second, female leaders in law enforcement have a direct influence on police op-

[4]Using the CSEW data from the Appendix Table published alongside the work by Flatley (2018), the author calculates the percent of the victims are female in the case of rape. The percentage of female victims of sexual offences is from Flatley (2018).

[5]Author's calculation.

erations to curb crimes against women. Distinctly, female legislators may have to negotiate with other legislators or make concessions on bills they introduce to prevent those crimes. Likewise, if law and order is not devolved to the local government, female representatives at the local level have to lobby national level decision-makers. Policewomen taking on leadership positions are much more likely to independently determine the policies, objectives, and structures of policing (Hewitt, 1991).

Could the increase in recorded female salient crimes be simple a measure of greater reporting willingness after appointing female chief officers? This could occur if the actual number of those crimes does not increase after the entry of women into top command positions, since such a rise can either be driven by an increase in the number of reported crimes or the actual number of crimes. The under-reporting of crime to the police is known to be particularly acute for female salient crimes, with many more offences committed than are reported to the police, becoming a major policy concern. Ministry of Justice (2013) admit that in England and Wales, only around 15% of those who experience sexual violence report to the police.[6] As a result, increased crime reporting is being considered a rule of thumb to identifying police responsiveness to those crimes (Soares, 2004b; Banerjee et al., 2012).

To disentangle the reporting effect from the change in actual crime incidences, this paper tests whether homicides against women that are less prone to reporting biases increase after appointing females as chief officers. The hypothesis that I advance and empirically establish is that women's access to police leadership positions leads to an increase in recorded female salient crimes is due not to a rise in actual crimes committed but, rather, to an increase in women's willingness to report.

I construct a dataset by combining detailed data sources on police workforce, crime and socio-economic characteristics at the police force area (PFA) level. Making use of this dataset, I find that the number of female chief officers appointed

[6]The reporting bias knows no geographical boundaries. which is also observed on an international stage (Skogan, 1976; Bachman, 1998; Muratore and Sabbadini, 2005; Joseph et al., 2017).

is strongly positively related to the number of recorded sexual offences and rapes, whereas it is strongly negatively associated to the number of homicides against women even after controlling for other potential important confounding factors, such as population, total police forces, and labour market conditions. These findings provide strong support for my hypothesis. In addition, I further control fixed effects for time-periods and PFAs respectively, and time trend across PFAs. Nevertheless, the results are maintained across various alternative specifications.

Even though I separate out actual changes in crime incidence from crime records, the above-mentioned approach captures not only changes in victim reporting behavior, but also changes in police recording practices that come from appointing women as chief officers, which could make interpreting results difficult. To differentiate between the two mechanisms, this paper utilises data on the number of transferred or cancelled records for sexual and rape offences, which could capture police behavior with respect to the removal of crime records, to investigate the association between the appointment of female chief officers and recording practices. The results suggest no significant association between police recording practices and women's appointments to police command positions. Nevertheless, there is a statistical power concern relating to interpretations of the results.

Another concern relates to other unobserved confounding factors such as budget reforms or institutional reforms that parallel the inclusion of women in leadership positions, which could have exerted a general influence on all forms of crime. To overcome the problem I employ data on gender-neutral crimes and repeat the same specifications. Consistent with the prediction, the results suggest that none of these crimes has any statistically significant association with the entry of policewomen into leadership positions.

To further tackle endogeneity concerns, apart from the ordinary least squares (OLS) approach, I also conduct an event study following Sun and Abraham (2021) to explore the causal impact of the first appointment of female chief officers on female salient crimes over the study period. Using variation in the timing of appointment of first female chief officers, I show that appointing female chief officers in particular PFAs increases documented female salient crimes and decreases the occurrence of violence against women in those PFAs.

My study relates also to other branches of the economics of crime literature focusing on sexual crimes. Card and Dahl (2011) find the significant effect of an upset loss by the local National Football League team on intimate partner violence. There are also some studies that highlight the effects of mandatory arrest laws (Iyengar, 2009) and no-drop policies (Aizer and Dal Bo, 2009) on reporting by domestic violence victims.[7]

Outside of the issue of sexual offences, this paper has a connection with a large literature regarding the effect of police on crime, where the police are often viewed as a primary factor generating a deterrence effect. Beginning with Levitt (1997), much of that literature focuses on IV strategies designed to overcome simultaneity bias, suggesting that the size of the police force significantly reduce crimes (Evans and Owens, 2007; Chalfin and McCrary, 2018). Leveraging a redeployment of police in response to a perceived terrorist threat, three studies also investigate the response of crime to changes in normal routines of policing (Di Tella and Schargrodsky, 2004; Klick and Tabarrok, 2005; Draca et al., 2011). There is a fast-growing strand of the literature investigating relationships between information technology and police productivity (Doleac, 2017). Unlike that literature, which focuses on interactions between police and criminals, this paper primarily studies the interaction between victims and police.

Finally, this work also echoes political theories of representative bureaucracy.[8] Most studies on representative bureaucracy is concerned with race and suggest that minority bureaucrats often implement policies to reduce the disparate treatment that minority clients have received historically (Meier and Stewart, 1992; Selden, 1997). Keiser et al. (2002) and Meier and Nicholson-Crotty (2006) take into account the contextual environment that influences whether sex will lead to active

[7] No-drop policies compel the prosecutor to continue with prosecution even if the victim expresses a desire to drop the charges and ceases to cooperate with the prosecution (Aizer and Dal Bo, 2009).

[8] The term "representative bureaucracy" comes from J. Donald Kingsley's book titled *Representative Bureaucracy*.

representation in the bureaucracy on this topic and find that increasing female representation in bureaucratic agencies helps women in the population.

The remainder of the article proceeds as follows. Section 2 provides a brief background on crime procedures in England and Wales and chief officer appointments. Section 3 describes the data. Section 4 offers the analysis of the appointment of female chief officers. In Section 5 I discuss my results. Section 6 concludes.

2.2 Background

In England and Wales, law enforcement is undertaken by police officers serving in 43 territorial police forces within one of those jurisdictions. These regional services are complemented by some specialist national agencies such as the National Crime Agency. All police forces are autonomous organisations, with responsibility for their own policing plan, resource allocation, budgets, and staffing decisions. On the other hand, police remuneration, minimum standards for police recruits, and legal knowledge examination have become increasingly centralised over time (Crawford and Disney, 2018). Figure 1 indicates the geographical boundaries of all territorial police forces.

2.2.1 Crime Procedures in England and Wales

The Home Office Counting Rules for Recorded Crime (HOCR) set out the manner in which police officers are to record crimes consistently and accurately. The HOCR specify that all reports of incidents, whether from victims, witnesses or third parties and whether crime related or not, will, unless immediately recorded as a crime, result in the registration of an auditable incident report by the police.

According to the National Crime Recording Standard, an incident will be recorded as a crime (notifiable offence) for 'victim related offences' if, on the balance of probability: (a) the circumstances of the victims report amount to a crime defined by law (the police will determine this, based on their knowledge of the law and counting rules); and (b) there is no credible evidence to the contrary immediately available. Nevertheless, recorded crimes may be later changed to transferred or cancelled records if they were determined that no notifiable crime occurred, recorded in error, or committed within the jurisdiction of another police force area.

When a crime is reported to the police, the prosecution process generally starts from the point. Depending on the type and seriousness of the offence committed, a decision whether or not to prosecute is made by the police service or the Crown Prosecution Service (CPS). In more serious or complex cases, prosecutors decide whether a person should be charged with a criminal offence. The CPS may advise police officers of the further evidence required before a charging decision can be made, or to strengthen the case to secure a successful outcome (Kaufman, 1980; Crown Prosecution Service, 2018).

2.2.2 Appointment of Chief Officers

Outside of London, each territorial police force is led by a team of chief officers headed by a chief constable, and supported by deputy chief constables, and assistant chief constables. While equivalent chief officers in the London MPS include the commissioner, deputy commissioner, assistant commissioners, deputy assistant commissioners and commanders. The role of chief officers is to provide strategic leadership for the force.[9]

Prior to 2012, the local police authority was responsible for hiring all chief officers, ensuring that there was local accountability for policing.[10] In contrast, all police officers below the rank of chief officer were appointed by, and under the direct control of, the Chief Constable. Every police force in England and Wales had a police authority. Despite the fact that some have more, most police authorities consisted of 17 members—9 councillors, 3 magistrates and 5 independents (Myhill et al., 2003). Among its members, a police authority shall at its annual meeting

[9]Police ranks below chief officers are listed as follows: constable, sergeant, inspector, chief inspector, superintendent, and chief superintendent.

[10]Mawby and Wright (2005) discuss the responsibilities of the local police authority and chief constable in detail.

appoint a chairman who plays a leadership role.

However, in 2012 police authorities outside London ceased to exist and were replaced by directly-elected Police and Crime Commissioners (PCCs) by way of the Police Reform and Social Responsibility Act 2011.[11] The PCCs have the legal power and duty to appoint, and where necessary, remove the Chief Constable. When it comes to appointing other chief officers (deputy chief constables and assistant chief constables), the Act states clearly that the chief constable must consult the PCC regarding the proposed appointment. Where differences in opinion occur regarding the proposed chief officer appointment these should be discussed and resolved locally between the PCC and Chief constable (College of Policing, 2018). Hence, the PCCs are actively involved with the appointment of deputy chief constables and assistant chief constables. In London the City of London Police continues to be overseen by a police authority, whilst the Mayor of London has responsibility for the governance of the London MPS. Similar arrangements to the London MPS also exist in the Greater Manchester Police since 2017.

2.3 Data

2.3.1 Crime Data

In England and Wales, police forces supply crime data to the Home Office, who are responsible for the collation and publication of crime data. The crime data used in this paper is drawn from the statistics of the Home Office. I obtain annual data on the number of crimes recorded at the police force area (PFA) level, which are available from 2002 through 2019 on a financial year basis.[12] As sexual and rape

[11] For the sake of brevity, PCC refers to Police and Crime Commissioner or Police Authority Chairman in this paper.

[12] There are 43 PFAs in England and Wales. However, the City of London PFA is a very small police force which covers the ''Square Mile'' of the City of London. This force is already excluded from all my analysis as it is not possible to obtain data on the covariates such as labor force participation for this area. This leaves me with 42 PFAs. They are Avon and Somerset, Bedfordshire, Cambridgeshire, Cheshire, Cleveland, Cumbria, Derbyshire, Devon and Cornwall, Dorset, Durham, Essex, Gloucestershire, Greater Manchester, Hampshire, Hertfordshire, Humberside, Kent, Lancashire, Leicestershire, Lincolnshire, London Metropolitan Police, Merseyside, Norfolk, Northamp-

offences are not comparable due to recategorization in the publications of crime statistics before and after 2005, my analysis for those offences starts from 2005. The dataset containing information on homicides against women is directly provided by the Home Office upon request, because the data in the publication version only include the aggregate count of homicides. Figure 2 shows the average number of sexual offences, rapes, and homicides against women in police force areas over the study period, respectively.

To investigate the effect of female chief officers on crime outcomes, I also obtain data on annual prosecution and convictions volumes of rape offences that are available from 2006. The recording system for these crimes is consistent over the period. I also collect the data on crimes that are not gender-specific, such as homicide, assault with injury, burglary, and motor theft. In order to investigate the effect on police recording practices, I collect data on the annual number of records for sexual offences and rapes that are transferred or cancelled. Specifically, I obtain the data from 14 police forces hold records from 2005 onwards. In terms of the rest of PFAs, I employ supplementary data from the Home office that has started to release the data from 2011.

2.3.2 Police Workforce Data

I also collect police workforce data at the PFA level from the national statistics of the Home Office for the period 2005-2019, which are originally sourced from police forces' personnel records. Forces keep records of all staff employed by the force on their human resource and payroll systems, as employees will not be paid if they are not on these systems. The key independent variable in this paper is the number

tonshire, Northumbria, North Yorkshire, Nottinghamshire, South Yorkshire, Staffordshire, Suffolk, Surrey, Sussex, Thames Valley, Warwickshire, West Mercia, West Midlands, West Yorkshire, Wiltshire, Dyfed-Powys, Gwent, North Wales and South Wales. One may wonder whether the inclusion of London Metropolitan Police Service (MPS) is biasing my results, since it is arguably very different from other police forces. For instance, police pay scale applies across England and Wales, yet officers in the London MPS are entitled to additional 'London weighting' and 'London allowance'(Crawford and Disney, 2018). Unlike other police forces, it has certain national responsibilities such as leading counter-terrorism policing and the protection of the royal family and senior members in the government. As a robustness exercise, I repeat the main regressions by excluding the London MPS.

of female chief officers appointed within a year, which comes from the joiners data including the number of new recruits to a particular police force by rank and gender.[13] Figure 2 shows the average number of female chief officers appointed from 2002 to 2019. Panels A and B present positive and strong correlations between female chief officers appointed and recorded sexual crimes and rapes, while Panel C presents a negative and strong correlation between female chief officers appointed and homicides against women.

To account for possible endogeneity of women's access to chief officer positions and control for potential confounding factors, I collect further data on police workforce. These consist of police employment data, such as police officers by gender and rank, as at 31 March each year, and police turnover data, such as number of officer leavers by rank and gender within one year.

Data on area characteristics are mainly obtained from the Annual Population Survey (APS) and the Annual Survey of Hours and Earnings (ASHE). The APS has the largest coverage of any household survey and enables the generation of statistics for local authorities in the UK. I aggregate labour market data drawn from the APS up to PFA level from local authority level. As the information on annual earnings, I extract local area mean earnings from the ASHE data that are the most comprehensive sources of earnings data in England and Wales, and then match them into police areas using local authority employment weights. Population figures by PFAs used in my analysis are extracted from publications by the Home Office.

[13]There are missing years for some forces. They are Thames Valley in 2006 and 2009; Suffolk from 2016 onwards; London MPS in 2017 and 2018; Dorset between 2005 to 2015.

2.4 Analysis of the Appointment of Female Chief Officers

2.4.1 Baseline Empirical Model

Table 1 provides the summary statistics for the main variables used in my analysis. To gauge the effects of appointing females to the role of chief officers on crimes, and particularly female salient crimes, I begin with a simple OLS model, which is specified as follows:

$$Crime_{at} = \alpha + \beta Female_{at} + \varphi X_{at} + \rho_a + \gamma_t + \varepsilon_{at} \tag{1}$$

Where $Crime_{at}$ represents the recorded crime for the following crime types: sexual offences, rape offences, homicide, assault with injury, burglary, and motor theft, in PFA a in year t. My study primarily focuses on sexual offences, rapes, and homicides against women but I examine gender-neutral crimes to identify whether other mechanism are in operations which have generalized effects on crimes. Observed crime counts are annual totals, but the observed police numbers are a snapshot as of 31 March. Thus, the number of the newly recruited female chief officers in that locality and year, $Female_{at}$, is used as the main explanatory variable. A set of time-varying control variables such as the number of male chief officers recruited, chief officers who leave from office by gender, total chief officers by gender at the beginning of the financial year, female officers in low-ranking positions joining and leaving, and total low-ranking officers, population size, unemployment by gender, and earnings by gender. ρ_a is area fixed effects used to absorb all area characteristics that are persistent over time. I also include γ_t, year fixed effects to take into account for non-parametrically adjust for national trend in crimes. ε_{at} is idiosyncratic error term. All standard errors are clustered at the PFA level to allow for possible correlated shocks to PFA-level crimes over time. Observations are also weighted by population size.

2.4.2 Baseline Results

2.4.2.1 Female Salient Crimes

Initially, I focus on female salient crimes and I present OLS results using variations of the model in equation (1). Table 2 demonstrates the impact of the addition of females to the seated chief officers on recorded female salient crimes. Considering the differences in severity levels across female salient crimes, I carry out an analysis of sexual offences and rape offences where the former represent those with different ranges of severity, and the latter stand for those with a high range of severity. In the first three columns the dependent variable is the logged number of sexual offences recorded, while in the last three columns that is the logged number of rape offences recorded. The coefficients on the number of newly added female chief officers, β, are positive and statistically significant in all specifications. To be more specific, OLS estimates for the key independent variable is statistically significant without control variables in columns (1) and (4). The size and statistical significance of them are little affected by the inclusion of many control variables in columns (2) and (4), and control variables plus area specific time trends in columns (3) and (6). Across specifications, I include year fixed effects and area fixed effects as well.

In columns (3) and (6), the results from my preferred specification with all controls are shown, suggesting that an increase in the number of newly appointed female chief officers is significantly associated with the number of recorded sexual offences and rapes. Specifically, appointing one female chief officers is strongly and significantly associated with a 4 percent increase in recorded sexual crimes and a 3 percent increase in recorded rapes. It is important to note that these estimates remain significant even when I control for the number of female officers in low-ranking positions. This result is valuable because it reveals that women being represented in more influential positions at the high level can directly affect female salient crimes, independent of large scale representation at the lower level in law enforcement.

As discussed above, the strong and positive association between female chief officers and female salient crimes could be attributed to more incidence, more recording, and/or more reporting. The next section tests whether the appointment of female chief officers leads to more incidence of violence against women.

2.4.2.2 Homicides against Women

Given the challenge of disentangling an increase in recording/reporting incentives from an increase in actual incidences of female salient crimes, this section focuses on examining the effect of appointing female chief officers on homicides against women where endogenous reporting is unlikely to be a concern. If appointing female chief officers leads to an surge in actual incidences of crimes against women, then greater recording/reporting incentives cannot be considered independent drivers of the results displayed in Table 2.

Using the data on homicides against women back to 2002, I run the same specifications as the analysis earlier. Table 3 reports the corresponding results for homicides against women. The results show that adding female chief officers is statistically significant and negative associated with homicides against women: 1 additional female chief officers is strongly associated with a decrease in homicides by 1 homicide. It is economically meaningful, when compared to mean value of around 5 homicides against women. This finding itself is important, implying that women's representation in the leadership of law enforcement agencies matters for women's safety. Furthermore, my result offers important evidence that an increase in recorded sexual crimes after a police force appoints more female chief officers is good news, driven by greater recording/reporting rather than greater incidence of sexual crimes.

Although I estimate the model for homicides against women in levels, without applying logarithmic transformations, because of its relative infrequency, as shown in Appendix Table A2.1, the results are robust to logarithmic transformations. Another way to corroborate my results is to exclude the London MPS from the regression estimates. It should be noted that the overall operational leader of the London MPS is the Commissioner who is appointed by Royal Warrant based on the recommendation of the secretary of state. Thus, it is entirely different for police forces in London and the rest of the country to run their process of appointing chief officers.[14] Given these differences, I further test whether the inclusion of the London MPS is biasing my results by excluding the London MPS. Appendix Table A2.2 offer the results from the repeated regressions. The results remain consistent and are robust to the exclusion of the London MPS, which increase my confidence in the validity of my results.

2.4.2.3 Police Recording Practices

If police recording practices improve in tandem with the rising number of recorded female salient crimes, it would be difficult to interpret an increase in those crimes as evidence in favor of a greater reporting willingness. Many studies have shown that there could be considerable distance remains between citizen "reported" and police recorded crime (Skogan, 1975). In this section, I use the number of transferred or cancelled records as a proxy for police recording practices, because it captures the quality of police action when recording crimes.

I examine the effect of the appointment of female chief officers on the change in transferred or cancelled records for female salient crimes, in order to get a better understanding of the extent to which I can attribute an increase in recorded female salient crimes to the change in victim reporting behavior. Table 4 displays these results for sexual offences and rape offences, respectively. The estimates across columns are insignificant at conventional levels, implying that police recording practices might not be affected by the appointment of female chief officers. However, as I discuss in detail later, these findings remain inconclusive, owing to a statistical power challenge.

[14] Also, in terms of responsibilities, unlike other police forces, the London MPS has unique responsibilities with respect to leading counter-terrorism policing. As a result, chief officers in the London MPS may be more sensitive to terrorist threats rather than other crime types.

2.4.2.4 Gender-Neutral Crimes

One may wonder whether the addition of female chief officers plays a role in other crime types, apart from female salient crimes too. This could occur if women are less competent than their male counterparts at policing, and thus greater female representation in the influential positions in law enforcement leads to a general slump of law and order.[15]

Table 5 offers estimates of the influences of the entry of female chief officers on the logged number of gender-neutral crimes: homicide, assault with injury, burglary, and motor theft, respectively.[16] None of these crime categories shows any significant connection with the change in the number of female chief officers appointed. This helps to show that my main result may not be driven by a spurious correlation or unobservable factors that might affect crime at large. Nevertheless, I will present power analysis below and the interpretation of null effects should be cautious.

2.4.2.5 Power

To disentangle the victims' reporting mechanism from police recording mechanism and general mechanism, I have examined the effect of appointing female chief officers on transferred/cancelled records and other crime records, and found statistically insignificant effects. These results may not suggest that the two mechanisms are not in operation, because it can be difficult to distinguish between lack of effect and lack of power. This difficulty suggests an innovation in reporting results. For each null finding, I compute the minimum detectable effect size (MDE),[17] that is, the effect that would have been detectable within 80% power at the 5% significance level ex post. In other words, the MDE is the effect size below which I may not be able to distinguish that the effect is different from zero, even if it is.

Results of power calculations and corresponding estimated effect size are pre-

[15] The role of female officers has long been controversial (Martin and Jurik, 2006).

[16] As there are several observations that equal to 0, I use log(homicide+1) as the outcome variable. Nevertheless, the results are robust when excluding those observations and using log(homicide).

[17] MDE are calculated with Stata's power twomeans command.

sented in Table 6. Given my current sample size, setting $\alpha = 0.05$ and $1 - \beta = 0.8$, I would be able to detect an effect of 40.90 for transferred/cancelled sexual offences, which is equivalent to 33 percent of the mean value, and 0.27 standard deviations of the records transferred/cancelled. For transferred/cancelled rape offences, the MDE is 19.33, which is equivalent to 36% of mean value, and 0.28 standard deviations. They show that I am not powered to detect very small effects, while I could expect to detect moderate and large effects. As MDE are larger than estimated effect sizes, I cannot rule out the police recording mechanism with confidence.

Panel B displays the estimates and MDE for other crime outcomes. Relative to the estimated effect sizes, MDE are larger, and imply a large risk of not detecting any smaller effects. Nevertheless, they are ranging between 0.02 and 0.06 percent of mean log units. In addition, none of the point estimates show effect estimates in excess of 0.01 percent of the mean log values, suggesting that the effects are economically small. Therefore, I can say the appointment of female chiefs does not have meaningfully large impacts on other crime types. Overall, I may be underpowered to detect the police recording and general mechanisms, and those null effects should therefore not be overinterpreted.

I summarize my key results as follows. On the one hand, appointing female chief officers significantly increases the number of recorded female salient crimes. On the other hand, evidence shows that the actual incidence of violence against women decrease. Overall, my results indicate that increasing essential inputs into the production of services in criminal justice system after appointing female chiefs may decrease violence against women.

2.4.3 Event Study

2.4.3.1 Endogeneity concerns

In equation (1), my coefficient of interest is β which measures the impact of adding females to the chief officer pool on the number of female salient crimes. However, for β to identify the causal effect, I have to address potential concerns that the entry of women into leadership positions is correlated with some confounding factors correlated with female salient crimes. For instance, one may be concerned that more female chief officers leaving from office are correlated with both appointing women to the positions of chief officer and female salient crimes, and my estimates are biased. I account directly for this by controlling for the number of female chief officers who leave the service. It could be also the case that police forces make decisions whether to hire additional female chief officers based on the number of existing chief officers, which subsequently affect crimes. Considering diversity, police forces that have fewer females in senior management team may hire more females. To account for this, I also control for the number of female chief officers at the beginning of the year.

It is possible that the addition of females to the top positions is accompanied by a rise in female frontline officers. Thus, the estimates might not be clean from biases due to an increasing in female representation across ranks. To address this issue, I also include the number of appointed female officers in low-ranking positions as a control for varying diversity across ranks of the police, though this in fact is a potential channel through which adding female chief officers affect crimes. Furthermore, I controls for other factors that the literature identifies as important determinants of crimes: population size, unemployment rate by gender, and earnings. In addition to including area fixed effects and year fixed effects, I also control for area specific time trends.

Reverse causality may be another concern which biases my OLS estimates, as it is noticeable that areas with greater number of female salient crimes tend to hire more female chief officers. Even though I control for potential confounding factors, employing various measures of female salient crimes, and using additional data sets, in terms of identifying causality, elaborating a suitable event study is still of necessity.

2.4.3.2 Event Study Specification

To overcome the endogeneity issue, I follow Sun and Abraham (2021), who develop DiD (difference-in-difference) estimators of the cohort-and-period specific effects to estimate dynamic effects in setting where there is variation in treatment across units. Their methodology is robust to treatment effect heterogeneity which recently has gathered a lot of attention.[18] As each police force has appointed at least one female chief officer, my analysis estimates separate event studies for each treatment cohort (defined by the year the first female chief officer was appointed) over the study period.[19] Based on their methodology, I estimate the following equation:

$$Y_{at} = \sum_{e=1}^{E} \sum_{\substack{k=-5 \\ k \neq -1}}^{5} \beta_{e,k}(\mathbb{1}\{E_a = e\} FirstFemaleChief^{k}_{at}) + \sigma_a + \phi_t + \varepsilon_{at} \tag{2}$$

Where time windows span periods of one year each. Particularly, $\pm k$ ranges from 5 and -5 to respectively 1 and -2 years, as -1 is omitted. Each lag(lead) takes the value of the main regressor -$k(k)$ years away from the first appointment of female chief officer. $FirstFemaleChief^{k}_{at}$ is a set of relative event-time dummies, that take value of 1 if time t is k years after (or before, if k is negative) the first female chief officer appointment. e represents cohorts, different time periods in which police forces appoint female chief officers. Each estimated parameter is a weighted average of the impact of the appointment k years away (i.e., pre or post). The coefficients of interest $\beta_{e,k}$ can be interpreted as an average effect of the treatment on the treated periods after initial treatment.

[18] Recently, a growing literature has investigated the biases of the two-way fixed effects DD estimator associated with heterogeneous treatment effects (Sun and Abraham, 2021; Callaway and Sant'Anna, 2021; Athey and Imbens, 2021; Goodman-Bacon, 2021). In particular, such an estimator equals a weighted average of all possible simple 2x2 DDs that compare one group that changes treatment status to another group that does not.

[19] Their procedure ensures nonnegative weights and better sheds light on dynamic treatment effects.

2.4.3.3 Event Study Evidence

In this section I move to estimate the impact of appointing female chief officers by estimating equation 2. I start by presenting evidence on the effects on recorded female salient crimes. I later assess the impact on homicides against women, which is followed by the impact on crime outcomes.

The event study estimates will yield a biased estimate of the causal effect of appointing female chief officers if earlier appointment of female chiefs disproportionately occur in police forces that would have experienced a change in the female salient crimes even absent the appointment. Although I cannot rule out such a possibility, I can check for differential clearance trends between earlier and later appointment police forces in the years leading up to the appointment of female leaders.

Figure 3 displays the corresponding estimates for recorded sexual offences. Prior to appointing female chief officers, I find little evidence of differential group trends in sexual offences. For $k < 0$, all treatment coefficients never reach statistical significance, supporting the identification assumption. The effect occurs immediately after the appointment. Following the appointment event, the recorded sexual crime increases significantly by 6 percent. But the effects subsequently reach insignificance and appear to disappear.

Two factors may explain the pattern of effects. First, victims respond to news coverage of the appointment of female chief officers. That is, the detected reporting effect may depend on media coverage of such appointments. A large effect immediately in the first year after treatment could suggest that female chiefs are more likely to draw the media's focus at the beginning of the appointment. Second, it might relate to police turnovers. Particularly, women police officers have much higher turnover rates than their male counterparts (Miller et al., 2017).

Figure 4 presents the corresponding estimates for recorded rape offences. For posttreatment periods, consistent with the pattern of sexual offences, recorded rape offences increase significantly by 9 percent in the first year of female chief officer appointment, and then the effects dissipate. However, unlike the above outcome I have examined, there is evidence of differential pretrends in rape offences. The existence of the pretrends makes me reluctant to interpret the estimated drastic increase in recorded rape counts too strongly.

Turning to the actual crime incidence, Figure 5 plots the coefficients from regressions that replace sexual offences with homicides against women for lead and lag indicators. I do not detect differential pretrends for homicides against women. However, police forces appointing female chief officers would experience significantly less violence against women immediately following the appointment. This is again consistent with the sexual crime figures and suggests that the appointment of female chief officers bring about more reporting and lead to less violence.

2.4.4 Female Chief Officers Who Stay

As discussed above, the short-lived effect of the appointment of female chief officers could be explained by high turnover. In this section, I examine the effect of the appointment of first female chief officers who stay after appointment by excluding police forces areas where the first female chiefs left the areas one or two years after the appointment.[20] In doing so, I am able to investigate whether female chief officer turnover is affecting the pattern of effects.

The same exercises as the previous section are repeated focusing on female chief officers who stay. Figures 6-8 report the results for sexual offences, rapes, and homicides against women, respectively, showing that the effects are not lasting. These estimates, both in terms of magnitude and significance, are strikingly similar to those for all appointments of first female chiefs. As a result, high turnover is unlikely to be a driver of the detected effect.

In a nutshell, the results of OLS regressions and event study analyses paint a very consistent picture, implying that female chief officers have meaningful effects on reducing violence against women. Considering the behavior of the three types

[20]The results are robust to excluding police forces areas where the first female chiefs left the areas one year after the appointment only.

of interacting agents—–police, victims, and criminals, I find a strong and significant elasticity for reporting/recording incentives with respect to women's access to police leadership positions.

2.5 Discussion

The main argument of my analysis is that victims, police and potential criminals respond to additional female voices in law enforcement leadership. In particular, I find that the more female chief officers police forces appoint, the more female salient crimes the police record. More importantly, I find that the actual incidence of female salient crimes declines after policewomen land the coveted leadership positions. The combination of these findings implies that women's access to police leadership positions substantially strongly positively affects essential inputs into the production of law enforcement services. This result is important per se, and also to complement the prior literature on women helping women particularly in the previously male-only domains.

My results also offer a piece of evidence that the increase in recorded female salient crime is not a part of an overall jump in crimes after the appointment of female chief officers. The police image is often associated with aggressive behavior and physical strength with fast-paced violence. The debate centering on female officers' lack of physical strength and inability to maintain an authoritarian presence has been ongoing for decades.[21] If it is simply the case that women are worse suited to police tasks, then overall surge in crimes might occur. Yet, I find no significant impact of the entry of policewomen into top command positions on crimes not targeted toward women, such as homicide, violence with injury, burglary, and motor theft. That is why this result justifies the unique role of top policewomen in serving women, especially in victimization situations.

The closest papers to mine are the parallel studies by Iyer et al. (2012), Miller and Segal (2019), and Amaral et al. (2021). They mainly focus on the impact of im-

[21] see, for instance, Balkin (1988).

plementing policies pertaining to women's representation, such as political reservation, affirmative action, and women police stations on female salient crimes. However, the possibility of media-induced reporting effect has been ignored in the literature. The event study evidence in the study suggests a potential channel relating to media coverage of female leadership that might drive victims' responses.

2.6 Conclusion

This study uses the lens of gender to explore the role of chief officer preferences in access to justice for victims. I offer the first comprehensive analysis of the effects of increasing women's ascension to top positions in law enforcement on female salient crimes. Importantly, to overcome the problem of joint determination of crime and police staffing, I conduct the event study analysis of the appointment of the first female chief officers.

I find that women serving in positions of law enforcement leadership significantly increase documented sexual crimes while, this rise is good news, which is due not to a rise in actual crimes committed. A greater willingness of female victims to report crimes against them to the police could potentially explain this rise. However, I cannot rule out the alternative explanation of improved police recording practices after the appointment of female chief officers with confidence.

Overall, my study makes a contribution to the related literature as it is the first to investigate how important amplifying women's voices in law enforcement leadership is to increase essential inputs into the production of law enforcement services. It is to be hoped that future research will cast more light on the role of law enforcement leaders from racial, religious, or ethnic minority groups in crime fighting.

Fig. 1: Police Forces in England and Wales

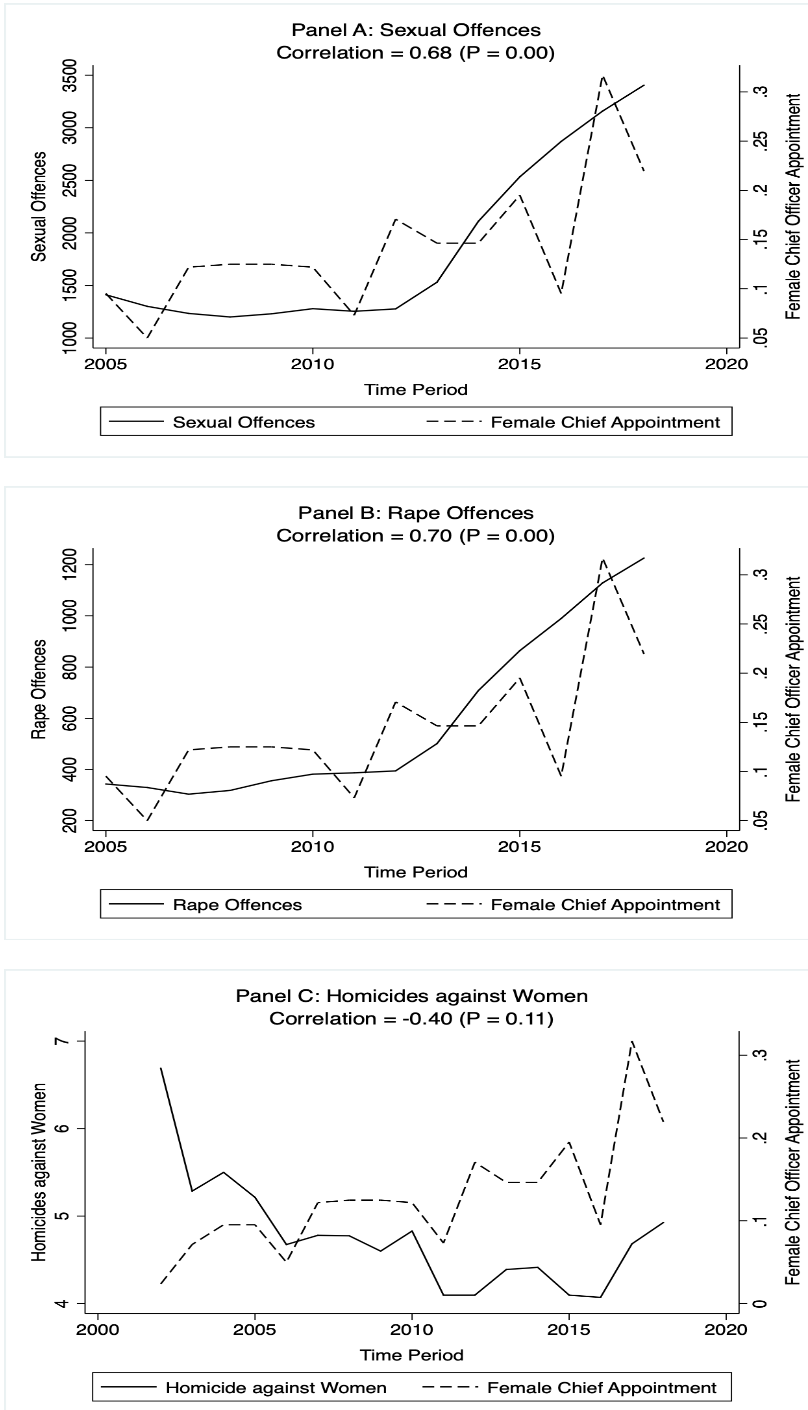


Fig. 2: Female Chief Appointment and Crime

Notes: The figure shows the plot of the appointment of female chief officers against the crime series. Panel A presents the average number of female chief officers appointed against the average number of sexual offence recorded, with correlation of 0.68. Panel B presents the average number of female chief officers appointed against the average number of rape offence recorded, with correlation of 0.70. Panel C presents the average number of female chief officers appointed against the average number of homicides against women, with correlation of 0.40.

Table 1: Summary Statistics

	Observations	Mean	SD	Min	Max
Panel A: Police Workforce					
Female Chief Officers Appointed	699	0.13	0.37	0	3
Male Chief Officers Appointed	699	0.43	0.72	0	8
Female Frontline Officers Appointed	699	53.65	90.61	0	986
Male Frontline Officers Appointed	699	113.75	224.35	0	2,375
Panel B: Crime records					
Sexual Offences	573	1,846.42	1,960.14	312	17,608
Rape Offences	573	589.54	691.47	67	6,421
Cancelled or Transferred Sexual Offences	402	120.82	146.57	0	1,021
Cancelled or Transferred Rape Offences	402	53.28	69.24	0	561
Prosecution Rape Offences	531	93.89	114.19	11	878
Conviction Rape Offences	531	55.47	59.39	6	462
Homicides against Women	699	4.78	6.01	0	54
Homicides	699	16.31	24.10	0	225
Violence with Injury	699	10,098.19	10,420.64	2,009	85,845
Burglary	699	13,013	14,999.92	1,267	113,427
Motor Theft	699	3,304.70	5,547.262	170	57,269
Panel C: Other Variables					
Female Chief Officer Leave	699	0.15	0.46	0	6
Male Chief Officer Leave	699	0.81	1.10	0	9
Female Chief Officers (31 March)	699	0.84	1.14	0	9
Male Chief Officers (31 March)	699	3.98	3.76	1	35
Female Frontline Officer Leave	699	33.67	43.25	1	478
Male Frontline Officer Leave	699	144.78	198.28	13	1674
Frontline Officers (31 March)	699	3,162.20	4,529.61	817	33,783
Population (000)	699	1,298.43	1,117.87	487.80	8,664.95
Female Unemployment Rate	699	5.87	1.73	2.15	11.75
Male Unemployment Rate	699	6.64	2.26	2.47	14.92
Female Earnings (£)	699	19,095.94	2,909.43	13,708.24	30,676.90
Male Earnings (£)	699	31,744.26	5,216.79	21824.00	53493.11

Notes: This table displays observations, means, standard deviations as well as the minimum and maximum values for variables of interest. Frontline officers represent officers with ranks below chief officers.

Table 2: OLS Estimates of the Appointment of Female Chief Officers on Sexual Offences and Rape Offences

	Sexual Offences			Rape Offences		
	(1)	(2)	(3)	(4)	(5)	(6)
Female Chiefs	0.035**	0.046**	0.039**	0.031**	0.042**	0.027**
	(0.013)	(0.017)	(0.015)	(0.012)	(0.018)	(0.012)
Year Fixed Effects	YES	YES	YES	YES	YES	YES
Area Fixed Effects	YES	YES	YES	YES	YES	YES
Controls	NO	YES	YES	NO	YES	YES
Area Time Trends	NO	NO	YES	NO	NO	YES
Areas	42	42	42	42	42	42
Observations	573	573	573	573	573	573

Notes: This table reports estimates by running OLS regressions. The sample consists of 42 police forces in England and Wales between 2005 to 2019. The dependent variables are logged number of sexual offences and rape offences recorded, while the key regressor is the number of women appointed as chief officers during a year. Controls include the number of male chief officers recruited, chief officers who leave from office by gender, total chief officers by gender at the beginning of the financial year, female officers in low-ranking positions joining and leaving, and total low-ranking officers, population size, unemployment by gender, and earnings by gender. All statistics are weighted by police force area population. Standard errors, clustered at the police force area level, are reported in parenthesis. Significance at the 1%, 5%, and 10% levels is indicated by ***, **, and *, respectively.

Table 3: OLS Estimates of the Appointment of Female Chief Officers on Homicides against Women

	Homicides against Women		
	(1)	(2)	(3)
Female Chiefs	-2.070***	-1.734***	-1.079**
	(0.692)	(0.613)	(0.518)
Year Fixed Effects	YES	YES	YES
Area Fixed Effects	YES	YES	YES
Controls	NO	YES	YES
Area Time Trends	NO	NO	YES
Areas	42	42	42
Observations	699	699	699

Notes: This table reports estimates by running OLS regressions. The sample consists of 42 police forces in England and Wales between 2002 to 2019. The dependent variables are the number of homicides against women, while the key regressor is the number of women appointed as chief officers during a year. Controls include the number of male chief officers recruited, chief officers who leave from office by gender, total chief officers by gender at the beginning of the financial year, female officers in low-ranking positions joining and leaving, and total low-ranking officers, population size, unemployment by gender, and earnings by gender. All statistics are weighted by police force area population. Standard errors, clustered at the police force area level, are reported in parenthesis. Significance at the 1%, 5%, and 10% levels is indicated by ***, **, and *, respectively.

Table 4: OLS Estimates of the Appointment of Female Chief Officers on Transferred/-Cancelled Sexual Offences and Rape Offences

	Sexual Offences			Rape Offences		
	(1)	(2)	(3)	(4)	(5)	(6)
Female Chiefs	-9.775	0.128	-5.643	13.024	-1.517	-7.612
	(11.873)	(14.759)	(15.912)	(12.166)	(6.699)	(8.520)
Year Fixed Effects	YES	YES	YES	YES	YES	YES
Area Fixed Effects	YES	YES	YES	YES	YES	YES
Controls	NO	YES	YES	NO	YES	YES
Area Time Trends	NO	NO	YES	NO	NO	YES
Areas	42	42	42	42	42	42
Observations	402	402	402	402	402	402

Notes: This table reports estimates by running OLS regressions. The sample consists of 42 police forces in England and Wales. The dependent variables are the number of sexual offences and rape offences that are transferred or cancelled, while the key regressor is the number of women appointed as chief officers during a year. Controls include the number of male chief officers recruited, chief officers who leave from office by gender, total chief officers by gender at the beginning of the financial year, female officers in low-ranking positions joining and leaving, and total low-ranking officers, population size, unemployment by gender, and earnings by gender. All statistics are weighted by police force area population. Standard errors, clustered at the police force area level, are reported in parenthesis. Significance at the 1%, 5%, and 10% levels is indicated by ***, **, and *, respectively.

Table 5: OLS Estimates of the Appointment of Female Chief Officers on Other Crime Types

	Homicides	Injury Violence	Burglary	Motor Theft
	(1)	(2)	(3)	(4)
Female Chiefs	-0.037	0.004	0.019	0.020
	(0.039)	(0.009)	(0.012)	(0.016)
Year Fixed Effects	YES	YES	YES	YES
Area Fixed Effects	YES	YES	YES	YES
Controls	YES	YES	YES	YES
Area Time Trends	YES	YES	YES	YES
Areas	42	42	42	42
Observations	699	699	699	699

Notes: This table reports estimates by running OLS regressions. The sample consists of 42 police forces in England and Wales between 2002 to 2019. The dependent variables are the logged number of homicides against women, violence with injury, burglary, and motor theft, respectively, while the key regressor is the number of women appointed as chief officers during a year. Controls include the number of male chief officers recruited, chief officers who leave from office by gender, total chief officers by gender at the beginning of the financial year, female officers in low-ranking positions joining and leaving, and total low-ranking officers, population size, unemployment by gender, and earnings by gender. All statistics are weighted by police force area population. Standard errors, clustered at the police force area level, are reported in parenthesis. Significance at the 1%, 5%, and 10% levels is indicated by ***, **, and *, respectively.

Table 6: Ex-Post Minimum Detectable Effect Sizes (MDE)

	(1)	(2)	(3)	(4)
	Panel A: Police Recording Practices			
	Transferred/Cancelled			
	Sexual	Rape		
Female Chiefs	-5.643	-7.612		
	(15.912)	(8.520)		
	[-37.78, 26.49]	[-24.82, 9.60]		
MDE	40.90	19.33		
Observations	402	402		
	Panel B: Other Crime Types			
	Homicides	Injury	Burglary	Motor Theft
Female Chiefs	-0.037	0.004	0.019	0.020
	(0.039)	(0.009)	(0.012)	(0.016)
	[-0.12, 0.04]	[-0.01, 0.02]	[-0.01, 0.04]	[-0.01, 0.05]
MDE	0.16	0.14	0.16	0.21
Observations	699	699	699	699

Notes: This table presents the estimated effect of female chiefs with standard errors in parenthesis and 95% confidence intervals in brackets, and calculated minimum detectable effect sizes (MDE). Panel A reports the MDE of transferred/cancelled sexual and rape offences compared to estimated effect sizes with $\alpha = 0.05$ and 0.8 power. Panel B reports the MDE of other crime types compared to estimated effect sizes with $\alpha = 0.05$ and 0.8 power. MDE are expressed in the same units as the outcome variables.

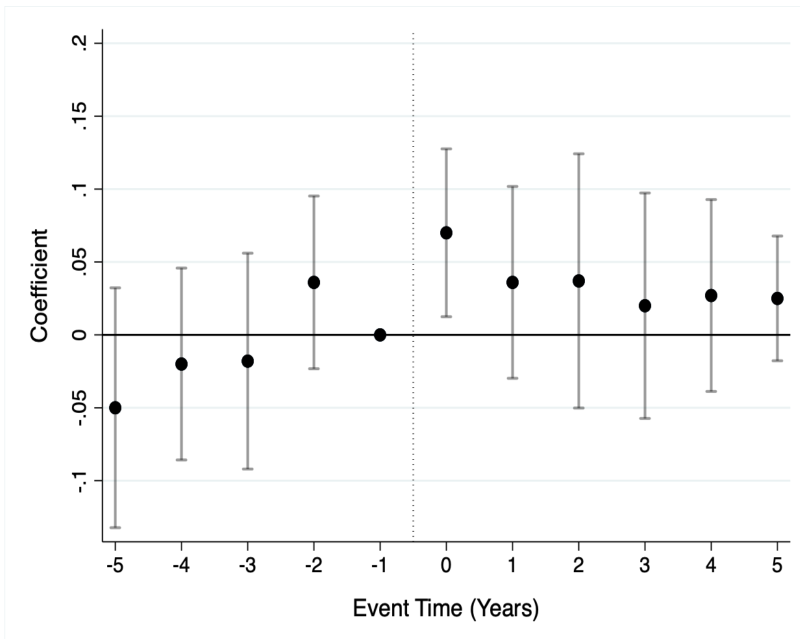


Fig. 3: Event Study Analysis for Sexual Offences

Notes: The figure presents the event-study plots of the effect of the appointment of the first female chief officer on sexual offences over the period, following Sun and Abraham (2021). Standard errors are clustered at the PFA level and observations are weighted by PFA population. The dotted vertical line represents the time of treatment. Bands indicate 90 percent confidence intervals.

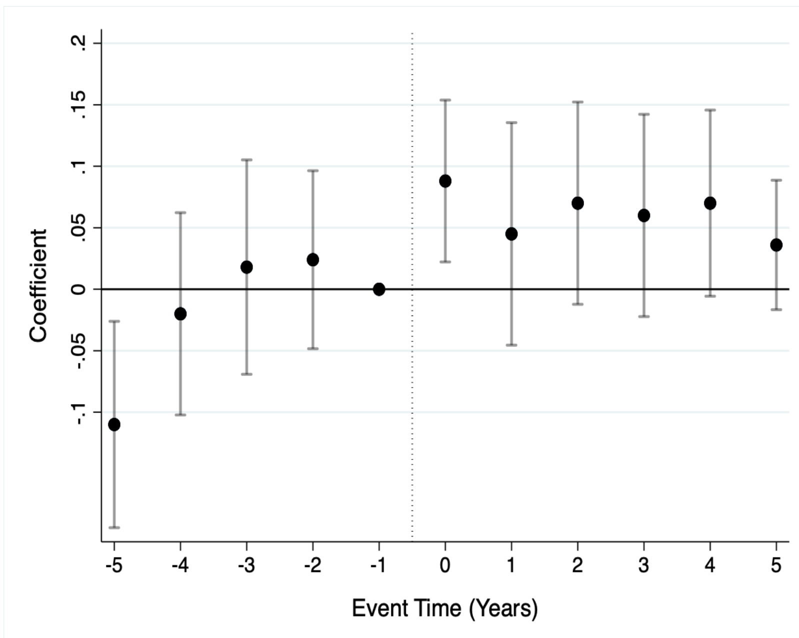


Fig. 4: Event Study Analysis for Rape Offences

Notes: The figure presents the event-study plots of the effect of the appointment of the first female chief officer on rape offences over the period, following Sun and Abraham (2021). Standard errors are clustered at the PFA level and observations are weighted by PFA population. The dotted vertical line represents the time of treatment. Bands indicate 90 percent confidence intervals.

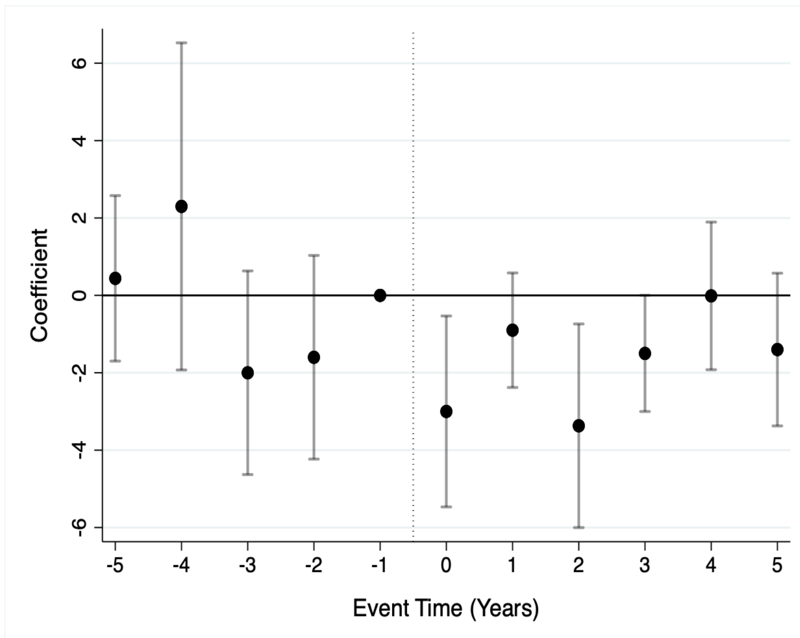


Fig. 5: Event Study Analysis for Homicides against Women

Notes: The figure presents the event-study plots of the effect of the appointment of the first female chief officer on homicides against women over the period, following Sun and Abraham (2021). Standard errors are clustered at the PFA level and observations are weighted by PFA population. The dotted vertical line represents the time of treatment. Bands indicate 90 percent confidence intervals.

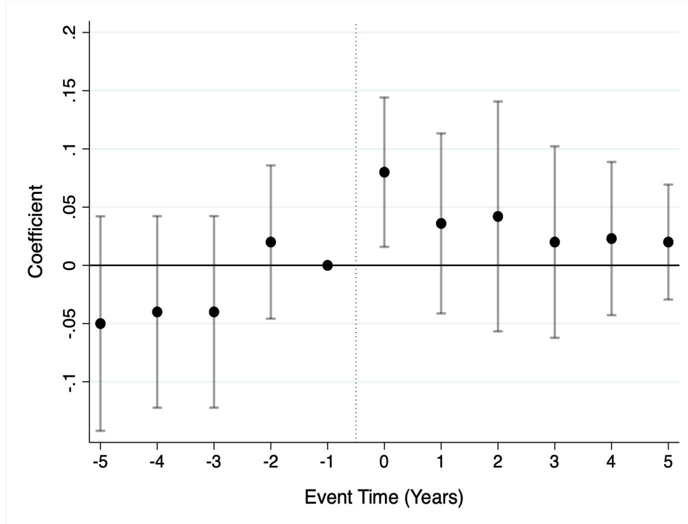


Fig. 6: Event Study Analysis for Sexual Offences (Female Chiefs Who Stay)

Notes: The figure presents the event-study plots of the effect of the appointment of the first female chief officer who stay on sexual offences over the period, following Sun and Abraham (2021). Standard errors are clustered at the PFA level and observations are weighted by PFA population. The dotted vertical line represents the time of treatment. Bands indicate 90 percent confidence intervals.

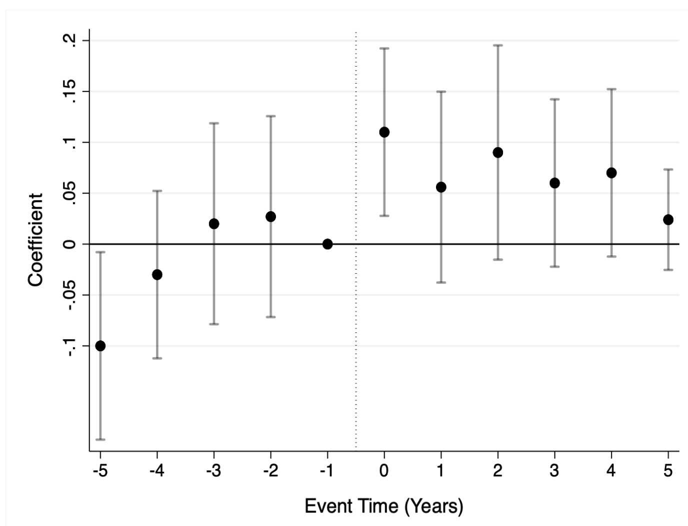


Fig. 7: Event Study Analysis for Rape Offences (Female Chiefs Who Stay)

Notes: The figure presents the event-study plots of the effect of the appointment of the first female chief officer who stay on rape offences over the period, following Sun and Abraham (2021). Standard errors are clustered at the PFA level and observations are weighted by PFA population. The dotted vertical line represents the time of treatment. Bands indicate 90 percent confidence intervals.

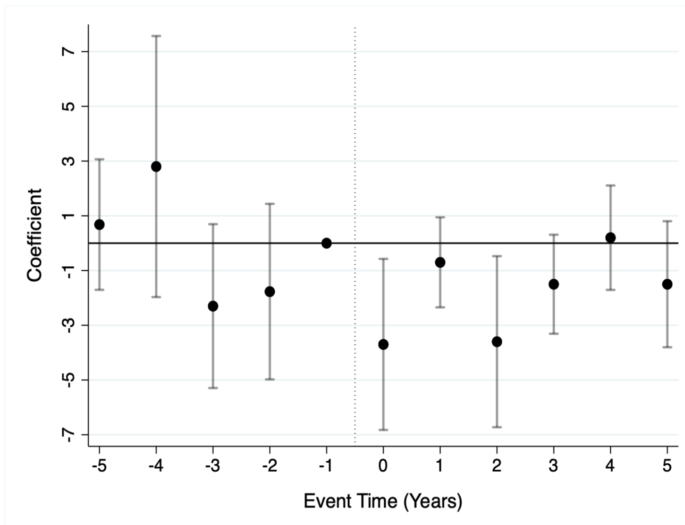


Fig. 8: Event Study Analysis for Homicides against Women (Female Chiefs Who Stay)

Notes: The figure presents the event-study plots of the effect of the appointment of the first female chief officer who stay on homicides against women over the period, following Sun and Abraham (2021). Standard errors are clustered at the PFA level and observations are weighted by PFA population. The dotted vertical line represents the time of treatment. Bands indicate 90 percent confidence intervals.

Table A2.1: OLS Estimates of the Appointment of Female Chief Officers on Homicides against Women: Robustness

	Log(Homicides against women)			Log(Homicides against Women + 1)		
	(1)	(2)	(3)	(4)	(5)	(6)
Female Chiefs	-0.091*	-0.132*	-0.122*	-0.072**	-0.099*	-0.087*
	(0.049)	(0.071)	(0.068)	(0.035)	(0.053)	(0.052)
Year Fixed Effects	YES	YES	YES	YES	YES	YES
Area Fixed Effects	YES	YES	YES	YES	YES	YES
Controls	NO	YES	YES	NO	YES	YES
Area Time Trends	NO	NO	YES	NO	NO	YES
Areas	42	42	42	42	42	42
Observations	649	649	649	699	699	699

Notes: This table reports estimates by running OLS regressions. The sample consists of 42 police forces in England and Wales between 2002 to 2019. The dependent variables are logged number of homicides against women without and with plus 1, respectively. The key regressor is the number of women appointed as chief officers during a year. Controls include the number of male chief officers recruited, chief officers who leave from office by gender, total chief officers by gender at the beginning of the financial year, female officers in low-ranking positions joining and leaving, and total low-ranking officers, population size, unemployment by gender, and earnings by gender. All statistics are weighted by police force area population. Standard errors, clustered at the police force area level, are reported in parenthesis. Significance at the 1%, 5%, and 10% levels is indicated by ***, **, and *, respectively.

Table A2.2: OLS Estimates of the Appointment of Female Chief Officers on Crime (excl. London MPS)

	Sexual Offences	Rape Offences	Homicides against Women
	(1)	(2)	(3)
Female Chiefs	0.043**	0.029*	-0.820*
	(0.018)	(0.017)	(0.435)
Year Fixed Effects	YES	YES	YES
Area Fixed Effects	YES	YES	YES
Controls	YES	YES	YES
Area Time Trends	YES	YES	YES
Areas	41	41	41
Observations	561	561	684

Notes: This table reports estimates by running OLS regressions. The sample consists of 41 police forces excluding the London MPS in England and Wales. The dependent variables are the number of sexual offences, rapes, and homicides against women, respectively. The key regressor is the number of women appointed as chief officers during a year. Controls include the number of male chief officers recruited, chief officers who leave from office by gender, total chief officers by gender at the beginning of the financial year, female officers in low-ranking positions joining and leaving, and total low-ranking officers, population size, unemployment by gender, and earnings by gender. All statistics are weighted by police force area population. Standard errors, clustered at the police force area level, are reported in parenthesis. Significance at the 1%, 5%, and 10% levels is indicated by ***, **, and *, respectively.

CHAPTER 3

Role Models Among Us: Experimental Evidence on Inspirations and Gender Disparities Set in Stones

Abstract: In this randomised controlled trial, we test the presence of such communications and their consequences from an inspirational lens. Focusing on a cohort of primary school students in India, we study the role modelling effect of historical statues. Students in the statues group were exposed to a short video of these locally present. The placebo group experienced an identical narrative on the same role models, with the visuals being composed of images of these role models and not their statue. There was a third pure control group. We find that a virtual tour to four statues of distinct role models in Jaipur leads to an increase in students' academic performance. Importantly, immediately after the intervention, students watching the statues video perform better in a memory test than the placebo and control groups. We postulate familiarity and susceptibility as the two fundamental constituents for relatability and as key mechanisms that fuel a motivational effect within the role modelling literature.

3.1 Introduction

Statues are erected as powerful pedagogical devices to transmit dominant political ideologies (Cohen, 1989; Kelly, 2015); or to serve as highly visible and respected symbols of prosperity attained by role models (Bell, 2008), and in the process paying them a tribute. Often they are associated with adding to the aesthetics of a city/-town, maintaining heritage of the people and attracting tourism too (Benhamou, 2020). However, it is not the complete story. Statues - as a mnemonic device - have also been identified to aid memory (Bower, 1970; Manuel, 2000) of a particular vision of 'ideal´ behaviour.

This is their role-modelling effect that channels information and influences behaviour by invoking past memories. For instance, portrait statues in ancient Roman society manifest the different roles that were then deemed appropriate for men and women in the society (Davies, 2008). Nevertheless, due to the lack of empirical data and methodological difficulties for comparative studies stemming from culture and context, very limited research has been performed on them, especially in economics. Using an experimental technique and limiting our attention on one city in Rajasthan, we assess these role-modelling effects in the short and medium run on children's academic performance. We also encounter the issue of gender disparity upon the analysis that is discussed in more detail later.

As a part of the research design, we exploit the effect that such statues of prominent individuals can have due to their accomplishments (by achieving high goals they serve as an inspiration to the society, or are at least expected to be by the installers)[1], visible grandeur[2], location (geographical accessibility) and by their origin and overall historical significance in the Indian society (being a part of the academic curriculum). In that sense, they cover all the tenets of role-modelling. For this reason, we label any treatment effects that we find as a result of role-modelling and introduce a conceptual framework of the underlying mechanisms. These further our understanding of the role-modelling literature and provide a completely

[1] Goal-attainability

[2] Aesthetics and being entertaining to look at or visit

new way to look at statues in policy discourses.

We start by identifying these role models and selecting a representative homogeneous sample of primary school children in urban parts of Jaipur - the study location. These students were then randomised at an individual level to three comparable groups at baseline.

Abiding with the school closures we resorted to an online experiment, benefiting largely by the support of the school staff. Without revealing the true intention of the study or our research questions, we conducted the experiment online using Qualtrics. To exercise control and ensure a somewhat neutral environment for the study, the baseline, intervention and endline phases were organised as a part of an activity during regular school periods. These were audited by a member of the research team and facilitated by the school teachers. Students filled online questionnaires in the baseline and watched the intervention videos right before filling those questions during the endline phase. The two waves of data collection happened in subsequent weeks.

This experiment is approved by the College of Social Sciences Research Ethics Committee of University of Glasgow (No: 400200239) and registered at the AEA RCT Registry (No: 0008055) before the intervention and survey data collection. The intervention consisted of a 6-minute video of four role-models[3] who had a statue in the city. We hired a production team to produce the short film on the role models and their statues. Our first (statues) group saw the videos on the statues - a virtual tour. The second (placebo) group saw a video on the same role models that was composed of their images. The third group served as a pure control. The videos had an identical narrative play in the background that was scripted carefully using the principles of the role-modelling and bearing close connections with the stories. Themes of hope, self-efficacy, hard-work, grit, breaking stereotypes and high aspirations were carefully embedded in these narratives.

A comparison between the two intervention groups with the control group, detected the effect of statues over and above the role modelling effect of these per-

[3] This was a compilation of a 90 seconds video for each.

sonalities by invoking additional feelings of familiarity and susceptibility. Using school's administrative data on academic performance we tracked the students for a period of one year and present effects immediately, one month and six months after the intervention.

We find that right after the intervention there was marked improvements in student's effort and career aspirations for public service. Students of the statues group had a 0.12 sd higher effect. This can be attributed to the additional memory stimulus received by only the statues group. Students of both the statues and placebo groups increased their aspirations by 0.14 sd.

One month after the intervention, students of the statues group increased their performance in Mathematics and Computer. On an average their scores increased by 0.21 sd and the number of students passing these exams by 0.23 sd. They also recorded an improvement of 0.04 sd in scores and 0.34 sd in passing the exam in social sciences, with no effects on languages (Hindi and English). All of these effects are strongly significant.

After 6 months, statues group students still performed better on Mathematics and Computer. The number of students passing the exam increased by 0.24 sd, which is significant at 5%. We detect no effects for the placebo group. The effect was driven by boys of the statues group. We find that boys scored 0.22 sd (immediate), 0.32 sd (1-month) and 0.25 sd (6-months) higher in the memory task, and mathemetics and computer exams. These effects are also strongly significant. These results and corresponding heterogeneities are discussed in more detail later.

Firstly, we contribute to a small but fast growing literature that studies role-modelling effects from movies and video clips on education attainment (Bhan, 2020; Riley, 2022), entrepreneurial aspirations (Bjorvatn et al., 2020; Dalton et al., 2021), labour market choices (Ahmed et al., 2022; McKelway, 2019) and youths' behaviour (Kearney and Levine, 2015). Yet, unlike that literature, we focus on the effect of real-life role models set in stones that have been existing for years. Instead of devising highly entertaining tools, we invoke susceptibility to otherwise invisibly camouflaging stone figures that may be appealing aesthetically, have a limited entertainment quotient. Inspiration may be lost in the crowd of concrete and buildings, but like a pair of car keys is readily present once you start looking.

Secondly, this study links to studies focusing on real-life role models as well. For instance, Kipchumba et al. (2021) show that visits from college students in randomly selected treatment schools impact students' aspirations. Porter and Serra (2020) investigate the effect of exposing students enrolled in introductory classes to successful and charismatic women who majored in economics at the same university. Much alike, Bettinger et al. (2018) and Chung (2000) find similar effects. The present study also looks into real-life role models made tangible due to their statues. We rely on recalling the content of students' memory rather than the direct interaction with role models. Here the principle mechanism is heightened sense of familiarity. We constantly draw comparisons from familiar and relevant others (WorldBank, 2014).

We see statues having an eventual role-modelling effect through the increased familiarity and susceptibility. Both of these are novel introductions to the role modelling literature which are the key components of relatability. Individuals' memory about role models may help to identify the common characteristics with those of the role models. The process may be viewed as communications between individuals with the role models, which is comparable with that of memory helping form common ground during communicative exchanges (Brown-Schmidt and Duff, 2016). Familiarity with and susceptibility to a given role model are the pre-conditions and determinants for how much an individual can relate to them. Role models rely on this tenet to 'speak with the viewer/audience/masses' and offer new information or shape their attitudes and behaviours. This conceptual deepening is illustrated empirically in our results. Using a placebo group that is equally familiar to role models, we manipulate the susceptibility and future familiarity of the statues group. Because the role models are locally presented in the form of statues, students might feel more relatable to them, which might help explain the lasting effects. For example, some ethnographic evidence from the northern India indicates that people set their reproductive goals based on the mortality experience of infants or children in their own locality (Lewis et al., 1958; Egerö and Hammarskjold, 1994). Nevertheless, it is only a hypothesis, as we do not measure the relatability in the data.

Thirdly, this study also contributes to the literature that examines the role modelling effect of political leaders on civilians' norm behaviour. Beaman et al. (2009, 2012) find that exposure to female political leaders increases women's aspiration and changes their gender norms. Also, conducting laboratory experiments, Gächter and Renner (2018) find that political leaders as role models can influence people's behavior in domains like charitable giving, tax evasion, and corruption. Our intervention videos include political leaders, enabling us to test for the effect of political leaders. We find that our intervention increase students' aspiration towards public services. Hence, this study may help explain why the Indian citizens are highly enthusiastic to be a part of public governance (IndiaGov, 2021; Forbesindia, 2020).

The rest of the paper is structured as follows. In section 2, we talk about the background literature on statues and their perception through a lens of role-modelling. Section 3 presents the research design with details about the intervention and the outcome variables. Section 4 includes our timeline, a description of the data and randomisation, along with the estimation strategy. Results are presented in section 5 and discussed in section 6 that also focuses on the potential mechanisms. Section 7 concludes.

3.2 Background

3.2.1 Statues, legacies and their implications

Statues (and monuments in general) built and installed to be seen and respected, often become invisibly blended into the daily hustle-bustles of people's lives (Marschall, 2017). People may barely pay conscious attention, let alone respect and actively appreciate these statues and their stories. Statues and monuments worldwide have recently become a matter of debate for what they represent, glorify and communicate in an unspoken manner.

Ever since the Roman imperial society[4], portrait statues were used to represent individuals worth emulation, and being regarded as 'role models' for 'the contemporary society' (Davies, 2008). While the former notion still prevails in today's societies (strategic remembrance), the finer latter caveat has been overlooked (selective forgetfulness) over the decades. While statues are erected in the present, often their installers do not take into account the preferences of contemporary society or contemplate the associated role-modelling effect. This has led to inconsistencies with what was deemed appropriate several hundred years ago, in the recent past and in the present. Also, completely ignored is the how they develop societies attitudes and behaviours.

3.2.2 Role models among us

Exposure to role models via multimedia channels can influence human behaviour and their aspirations (Bernard et al., 2014; Bhan, 2020; Chong and Ferrara, 2009; Chung, 2000; La Ferrara et al., 2012; Lafortune et al., 2018; Riley, 2022; Dalton et al., 2021). Role models have been found to affect household autonomy, improve academic performance, increase entrepreneurial spirit and savings, raise family incomes, appease gender biases and stereotypes, and foster hope and optimism across developing and developed countries. Interestingly there is plenty of evidence on their efficacy across different age-brackets too.

Studies with role models have widely used exposure to real life role models (in person and virtually) and fictitious characters (protagonists of films/television series/advertisements). Such interventions aim to offer a blend of information, encouragement and entertainment (in some cases) through an individual (or a group) that the subjects can associate with. Their achievements (often in the face of adversity) serve to instigate feelings of self-confidence, self-efficacy and hopefulness. A complementarity between beliefs about one's abilities and performance is central to their effect on different developmental outcomes. For instance, Bettinger

[4]Or even longer back in ancient Egypt. For more detail please see Blackman (1923).

et al. (2018) investigate how schools can increase students' perseverance in math by shaping students' beliefs in their abilities to learn.

Marking a departure from the standard literature, we explore whether 'statues' of real life historical icons can instigate a similar effect. In an attempt to comprehend the mechanisms of such an effect we extend our understanding of the underlying principles of role modelling. In that process, we present evidence on ways to strengthen role-modelling interventions. As mentioned formerly, public opinion and scientific discourses on the power and weaknesses of statues are mixed and limited. Through a role-modelling lens, we unearth their implications on psychology, academic performance and gender disparity in the following sections.

3.3 Research Design

Brocas and Carrillo (2020) argue that studies on children bring their own novel challenges for experimental designs that demand premeditated designs, and methodological adjustments based on the sample and context. Children grasp information, fuel their motivations, and adjust their behaviours in their own ways. For this reason, we adopt a host of measurement tools to collect information on their psychological characteristics and academic performance.

3.3.1 Sample

The sample consisted of 1571 students from eight schools in urban parts of Jaipur, India. They were selected based on a simple identification strategy that required each school to be co-ed (having both boys and girls), follow a Central Board of Secondary Education (CBSE) curriculum, have at least 40 students in class 4 and 5 each, and agree to share administrative data on school performance. We limit our attention on students in class 4 and 5 as younger children could lack the necessary comprehension and technical skills to partake in the intervention or the questionnaires. These students are randomised at an individual level to three homogeneous groups (one statues, one placebo, and one pure control). The schools were approached in April-May 2021 to gather consent.

Given the online nature of the research design and the ongoing remote schooling, information elicitation was purely digital. Nevertheless, we cross-checked the information collected from the surveys on children with the school's administrative records. Appendix TableA3.1 presents the summary statistics. Our average sample age at baseline is 9.5 years with 42 percent of the sample being girls and half of them in class 4 (and the other half in 5). Students scored at an average level of hope and aspiration levels with respect to the literature (Snyder, 2002; Bhan, 2020). At baseline, 75 percent of the students wanted to enter public services and the average sample scored 5.13 out of a 10 mark memory test. Exam score as a percentage is provided for five subjects including math, computer, English, Hindi, and social science. Their mean scores range from 0.82 to 0.86.

3.3.2 Intervention

As a part of the online RCT, we requested and received assistance from school teachers to ensure that participants pay attention to the videos. In August'21, all the teacher coordinators (1 teacher per class) were trained online to passively facilitate the intervention. During the intervention, students joined a zoom room shared by their corresponding teachers as they would on an average (at that time). When they joined, they were informed that they could participate in an activity and as a part of it, open a link that was shared in the chat box. This contained the link for a Qualtrics survey questionnaire.

Unlike the baseline[5], there were three distinct links for the endline. Contents of each link were treatment dependent, in so far that they first requested consent, followed by a video in the statues and placebo groups or no video in the pure control arm, and then survey instruments in the same order as the baseline. Hence, each

[5]Students received a common link during a similar zoom room organised and audited by the teachers. Since there was no treatment involved, each child received a common link that collected information on baseline covariates and outcomes of interest.

school organised three different zoom rooms, distinct to each group and carefully 'marked student attendance´ at the beginning and towards the end of the session. This step ensured (and to some extent prevented) compliance [6].

We hired a production company ('Braille Cam') to produce the intervention videos. Statues and placebo interventions comprise of a set of four role models, namely, Mahatma Gandhi, Indira Gandhi, Arjuna and Dr. B.R. Ambedkar. However, in the former, students watch a virtual tour of their statues in the city of Jaipur with a short narrative on their lives. Students in the latter watch a video with the same narrative and duration, but devoid of any statues and instead images of these role models stitched into a short video. The videos were narrated in Hindi, which is the local language of communication. The third group serves as a pure control. Each of the four videos was approximately a minute and half long and was delivered within one session in the local language (Hindi). Kids were exposed to the role-modelling content on individual screens[7].

The group watching the statues video (referred to as the statues group hereon) differed from the group watching the video composed of the images of these role models (referred to as the placebo group hereon) only in terms of the visual cues that they were exposed to. The two videos shared an identical narrative to specifically disentangle any treatment effects due to the virtual stimulus on children's memory. These role models were carefully selected from the twofold identification criteria. First, they should be a part of the children's academic curriculum and students may have some modest levels of familiarity to them. Second and importantly enough, they should have a statue in the urban part of Jaipur. This strategy enabled us after an exhaustive exercise to narrow down to four iconic figures in Indian history and mythology. These are Mahatma Gandhi, Dr. B.R. Ambedkar, Indira Gandhi, and Arjuna. Interestingly and much to our expected dismay, there was only one statue of a female potential role model.

[6] Students that had accidentally joined a group that did not correspond with their treatment status, were asked to leave and join the apt room before rolling out the link. For instance, if a student who belongs to statues group, entered the room for pure control students, s/he would be identified during the attendance and asked to join room for the statues group.

[7] As they are home and no households had siblings in the same class.

Mohandas Karmchand Gandhi, popularly known as Mahatma Gandhi worldwide, is a beacon of the strength of non-violence and peace. His video contains elements from his time in South Africa and India, from humble beginnings to becoming the father of the nation through sheer will and perseverance. The key message focuses on having faith in oneself - "The future depends on what you do today."

Dr. Bhimrao Ramji Ambedkar, was the Minister of Justice and Law and the head of the drafting committee of the Indian Constitution. This narrative is premised on hope, aspirations and grit. Overcoming discrimination and hopelessness at a very young age, B.R. Ambedkar secured a scholarship to earn a doctorate in Economics from Columbia University and became a social reformer. "They tried to bury me. They did not know that I was a seed."

Indira Gandhi, the iron lady of India, was the first female prime minister of independent India. Central to this video's theme are the emotions of confidence, courage and strength. These echo is a letter from her father"... Be brave, and all the rest follows. We work in the sun and in the light.".

Arjuna, third son of Pandu, is a popular character from Mahabharata (a Sanskrit epic from ancient India) famous for his archery skills. The video highlights his eagerness to learn and passion to perfect his skills. From practicing archery in the dark to focusing only on his target (The eye of a wooden fish) Arjuna excelled. Quoting Dronacharya[8] "... to achieve something, you must focus on it....and concentrate only on your target", the video emphasises the significance of goal-setting, focus and hard work.

The overarching theme of the intervention videos in general and the narrative to be precise, was to showcase the power of grit, hope and resilience in the face of adversity. With a focus on goals, consistent effort and power of will, each narrative delivers a key message, with the common element of 'you can do everything that you aim for´.

[8] Arjuna's teacher.

3.3.3 Main outcomes

We collected information on two main strands of outcomes of interest, namely psychological/motivational and academic performance. While the former is captured in the short-run using information elicited on a memory task and career aspirations, the latter is tracked over a span of 6 months since the intervention. To do so, we relied on administrative data from schools on exam performance of the students right before the intervention and one and six months afterwards.

Several instruments were employed to measure information on student motivation. Effort -as an objective proxy of motivation- was measured on a cognitive memory test Ariely et al. (2009) and an arduous spot the differences task. As a part of the online survey, students were displayed a sequence of five numbers for 10 seconds, which they had to memorise and re-write as an exercise in the subsequent page. This step was repeated thrice to ascribe a total score out of 10 for each student. In the second task, students were presented with two similar images with subtle differences in each. The students were asked to 'spot as many differences as they could between the two images and write them´ in an open text field in Qualtrics. However, the quality of data collected on this indicator is poor and subject to extensive scrutiny that has yet to be undertaken. Although, it is a novel indicator of effort, we had to drop it from the current analysis[9].

Information was collected on individual hope and aspirations. Alongside, we use information from children's career aspirations to create a dummy variable for whether they wish to go into public sector service, solely to capture the effect of the treatment on aspirations bearing in mind the role models and their accomplishments[10]. Information on these indicators was collected to identify any immediate boosts to motivation[11].

[9]Since there measure relied on an open field format for collecting the responses, students followed no consistent approach to answer. These range from answering just a final number to writing complete sentences for each of the total 15 difference.

[10]Mahatma Gandhi and Dr. B.R. Ambedkar excelled in academics (Law and Economics, respectively) before serving their nation in a not electoral-way. Indira Gandhi served the society by becoming a political and the first female prime minister of independent India.

[11]For information on other pre-registered but unanalysed indicators please refer to the PAP. These include a math test and word-completion task. As it was unexpected to identify any improvements

For collecting information on students' academic performance, we collected administrative data from schools on examinations conducted in March 2021, September 2021 and March 2022. Since different schools had by and large an identical curriculum, we compare the treatment effects on individual subjects and across STEM and languages (the STEM subjects include Mathematics and Computer, and the Language subjects include English and Hindi). In addition to assessing the treatment effects on scores, we also analyse the impact on the passing rates by discipline. In the short run, as a part of the survey questionnaire, we collect information on student's math and English language performance using simple math tests and word completion tasks, respectively[12]. In the former, students answer 10 simple arithmetic questions (suitable to their age group) and receive a score for each. In the latter they are asked to complete 5 incomplete words being rewarded with one point for each correct word formation.

Employing our research design and using these outcome variables, we answer three questions: (i) does exposure to statues motivate students; (ii) do they have a role modelling effect on academic performance; and (iii) can they explain disparities in academic performance in India? We answer these one by one in section 5 and 6.

3.4 Data and Estimation Strategy

Students of class 4 and 5 from the selected schools were organised to participate in an online session. Information on the contents of this session was withheld from them. These were organised and communicated to the students by the school teachers disguised as a regular school period. As the session started, the teachers took student attendance.

in these, we included them to ensure that the sample is balanced on measures of academic performance over and above the secondary data supplied by the schools.

[12]Besides the examination scores, we use these as useful proxies for learning levels at baseline for balancing checks. We do not report any analysis of treatment effects on these two variables as mid-way through the data collection we realised that it is unlikely that learning outcomes would change overnight. Since, these were administered right after the endline, it would have been much against the literature and intuition to observe any changes in these two.

These teachers were trained two weeks in advance and offered a script for that session. By training, scripting and anonymously auditing the session the research team ensured that no details about the experimental design, its intent or the outcome variables is communicated to the students. It was all the more difficult for the teachers to do so, as they or the school administration did not know of it themselves. Following Zizzo (2010), we adopted a non-deceptive obfuscation technique to avoid any potential experimenter demand effects. The RCT was presented as a general study to learn whether students like activity-based learning on online platforms. They then informed the students that it was an activity-based session and shared a link to an online survey questionnaire in the chat window. Teachers were informed that they should not help the students with any part of the survey and the students were instructed to not seek help from the teachers or their parents[13]. Endline surveys were conducted in a similar fashion with the exception of having groups segregated by the treatment assignments. Neither the teachers nor the students had any information on this randomisation.

3.4.1 Timeline

We administered a baseline survey in August 2021 to track the initial scores on the outcome variables, along with collecting essential information on other confounding factors (including but not limiting to age, gender, type of device used to attend online schools, class). The respondents to this questionnaire formed the participant pool and were randomised at an individual level.

In the next week, we rolled out the endline survey, of which, the intervention is the first part. We had the videos in-built as part of the online surveys. So any individual who wished to fill the endline survey - if belonging to the statues or placebo group - had to watch the videos in the beginning. Three types of endline surveys

[13] The students were also asked to keep their video on at all times to ensure the latter aspect. While a handful of parents intervened at a few occasions in baseline, it is safe to assume that randomisation balanced it across the different groups. Also at any such instance the teachers warned them instantly and that set example for others and avoided any repetitions of a similar malpractice. Additionally, by design, students were not allowed to chat amongst the participants of the call.

were administered to the participant pool, on the basis of their experiment assignments. As soon as the participants finished watching the videos[14], they progressed to the information elicitation tools for the variables of interest.

The online survey captured information on the outcome variables only. Information on student performance in school examinations both before and after the intervention were gathered using administrative records. Each of the two online sessions that included the surveys lasted for approximately 1 hour much alike a regular class would otherwise. The timeline is presented in Figure 1. Starting with a baseline survey in the last week, we conducted the intervention and the endline survey in the beginning of September. All the eight schools had their exams in March 2021, last week of September or early October 2021 and then again in March/April 2022. This allowed us to compare the academic performance of the students belonging to different groups, one- and six-months after the intervention.

After consulting local field workers and head teachers in the experiment schools, and looking into the media reports and government announcements, we also show the lockdown timeline with school status in Figure 1. In addition, using data from Google Mobility Reports (Google, 2022), we present the trends in mobility activities in parks and outdoor spaces in Jaipur in Figure 1.

As a precautionary measure, the Rajasthan government decided to close all schools on 14 March 2022 (The Hindu, 2020b). Later, in response to the first wave of covid-19 cases, the Rajasthan government announced a complete lockdown of the state on 21 March 2021 (The Hindu, 2020a). As shown in the mobility graph, the activity in the outdoor spaces collapsed after implementing the lockdown policy. The lockdown restrictions were partly released in May 2020, yet the schools remained closed (India Today, 2020). Since then, the outdoor activities started to recover.

In light of the delta variant rapidly spread during second wave of covid -19,

[14]To ensure attention over and above the presence of a school teacher to audit the session, we included some basic questions about each role model - and not their statue - as a part of the survey that followed the videos. Three questions appeared after each video that could not be replayed or recorded by the programming design. Only after answering them could a student proceed to the next video and repeat the exercise two more times.

the Rajasthan government imposed a state-wide lockdown in May 2021 (Business Standard, 2021). To secure livelihood of people, the government ease restriction measures on mobility in June 2021 (The Hindu, 2021). The mobility graph also shows that the outdoor activities increased accordingly. Because of school closures we resorted to an online approach towards data collection and intervention delivery. Our intervention was conducted at the start of Sep 2021. During the time, the mobility in outdoor spaces were not fully recovered, the schools remained closed, and all teaching activities were delivered online.

Importantly, Covid vaccination for kids was not approved and the Delta variant was leaving children vulnerable to Covid-19 (Rackimuthu et al., 2022; CNN, 2021b). Meanwhile, the reported cases of children contracting the virus increased, and experts suggested that parents should be more cautious in India (CNN, 2021a; Indianexpress, 2018). Therefore, statues and outdoor spaces were not inaccessible for children, yet school closure and the covid infection concerns may ebb regular exposure to these centrally located and frequently accessed statues by the students. This feature allowed our intervention to offer a jog down the memory lane for when the students would usually come across these statues in everyday life, while returning from school.

Due to the third wave of covid-19 cases, the lockdown was re-imposed in January 2022, while Rajasthan was completely unlocked and the schools were reponed after almost 2 years in February 2022 (Abp, 2022). As suggested in the mobility figure, the outdoor activities were fully recovered. We assume that this point on, a month before the examinations, students had an almost pre-covid level of exposure to these statues. However, only the statues group students may be now made more familiar to these iconic landmarks.

3.4.2 Randomisation and balance

Randomisation -at an individual level- was performed on Microsoft Excel, such that each participating student was assigned uniquely to one of the three groups -

statues, placebo or pure control. The statues group witnessed a virtual tour to four monumental statues of Jaipur with a short narrative in the background about the individual in that statue; the placebo group watched a video about the same role model that was prepared using images of those role models (but not their statue) with the same narrative as in the statues arm. The third group acted as a pure control. The design of experiment allocations is depicted in Figure 2.

As shown in Table 1, the three groups are comparable across a range of baseline characteristics. These include information on confounding factors (like age, gender, or the type of device used), psychological outcomes (hope, aspiration or memory task score) and academic performance parameters across different subjects. Column 7 presents the p-values of an F-test of mean equality across the three groups for these different variables. Since these are all insignificant, it is safe to say that the sample is balanced across the treatments. Additionally, as indicated in Appendix TableA3.5, attrition rate was also balanced across the three groups. Overall, we had approximately 27% attrition from baseline on an average, which is not surprisingly high for online experiments (Özler et al., 2021; Chin et al., 2021).

3.4.3 Empirical strategy

We estimate the treatment effects using an ANCOVA specification of the form:

$$Y_{it} = \beta_0 + \beta_1 Statues_i + \beta_2 Placebo_i + Y_{i0} + \varphi X'_{i0} + \varepsilon_i \tag{1}$$

where Y_{it} is the outcome of interest for student i in time t. $Statues_i$ is a binary indicator equal to one if student i is assigned to the statues group, $Placebo_i$ is a binary indicator equal to one if student i is assigned to the placebo group. Y_{i0} is the baseline value of the outcome. X'_{i0} is a vector of baseline characteristics for student i such as gender, age, class and among others. ε_i is the error term. Robust standard errors are calculated to allow for heteroskedasticity. β_1 and β_2 coefficients capture the effects of students watching a video of statues and a video of role models only, respectively.

As presented earlier, we try to ensure compliance in a number of ways. First, we have the intervention videos in-built as part of the online surveys and the videos cannot be speeded up or slowed down. As a result, participants have to watch the intervention videos before progressing to fill the endline survey. Second, school teachers carefully audited each session and marked the attendance throughout the session. Students that had accidentally joined the opposite of their treatment assignment, were asked to leave, and join the apt room. Third, we added some basic knowledge questions about each role model, as a part of the surveys that followed the videos. However, ensuring perfect compliance is difficult for our online sessions, because it is possible for students to be easily distracted by social media or other sites while videos are playing. Therefore, we are estimating the intention-to-treat (ITT) effect, that is an estimate of the effect of the program on those assigned to watch the videos, regardless of their take-up.

Gender disparities in education attainment have remained deep and persistent in India (World Economic Forum, 2020). We are also interested in identifying the effects by gender as well as the difference in effects between female and male students. Therefore, we estimate the following equation:

$$\begin{aligned} Y_{it} = \gamma_0 + \gamma_1 Statues_i + \gamma_2 Statues_i \times Female_i + \gamma_3 Placebo_i + \gamma_4 Placebo_i \times Female_i \\ + \gamma_5 Female_i + Y_{i0} + \phi X_{i0}^{'} + \epsilon_i \end{aligned} \tag{2}$$

where $Female_i$ is a binary indicator equal to one if student i is a female student and 0 otherwise. γ_1 measures the statues intervention on male students, while the sum of γ_1 and γ_2 measures the statues intervention on female students, and γ_2 measures the statues intervention on the gap in outcome between female and male students. Similarly, γ_3 measures the placebo intervention effect on male students, the sum of γ_3 and γ_4 the placebo intervention effect on female students and γ_4 the placebo intervention effect on the gender gap.

In the next section we present our findings for the effects. We assess these effects immediately, 1-month and 6-months after the intervention depending upon the outcome variable and the instrument.

3.5 Results

3.5.1 Main findings

A 6-minute treatment video consisting of four prominent Indian role-models had a striking effect on children's memory test. The video consisting of the statues only had an instant memory stimulus and heightened effort in a memory-based number retention task. Alongside, for either of the groups receiving a role-modelling treatment we find an increased aspiration towards public service. Academic performance of the statues group students increased a month after and the number of students scoring below the passing threshold fell. Effects on passing sustained after six months.

In Table 2, we report results from a survey questionnaire conducted right after the intervention. We find an increase of 0.12 sd in a memory task at 10 percent level of significance only for the statues group. Aspirations towards a public service oriented career increased for both the statues and placebo groups. The statues group recorded an increase of 0.14 sd with a slightly lower effect on the placebo group, both of which are significant at 5 percent[15].

Although the point estimates of the memory test suggest a substantially larger positive reaction to statues videos compared to the placebo videos, we cannot reject that the impacts are equal, with p-values of the test for equality of 0.31 and 0.29 for the specifications excluding and including controls, respectively in Table 2. Similarly, we cannot reject that the effects of statues videos on public service aspiration are the same as for placebo videos. These suggest that some caution is warranted in interpreting the short-term effects of comparing the statues videos with the placebo

[15] As students had the freedom to select an option of 'do not know' in the question for aspirations, we score it as missing. Hence, the analysed samples are slightly different across the two variables and in different tables.

videos on the memory stimulus and aspirations.

Table 3 shows the estimates for the impact of the intervention on the score and the average number of students passing the exam for an aggregated variable of the two STEM subjects (Computer and Mathematics). In Table 4, we show the effects on languages (Hindi and English) and social sciences. We report the effects both 1- and 6-months after in panel (a) and (b), respectively. In Table 3, we find that in the medium run the number of passing students in statues group increased by 0.23 sd (column 4) with the effect increasing to 0.24 sd after 6 months. Both of these effects are significant at 5%. As in column 2, statues group students' score increased by 0.21 sd with the effect being strongly significant.

Turning to the comparison with the placebo videos in the medium run, we cannot reject that placebo videos have zero effect on score of math and computer, but neither can we reject that its effect is equal to the statues one, though the impact estimate of the placebo is much smaller compared to that of the statues in columns 1 and 2 of Table 3. However, as shown in columns 3 and 4 of panel (a), we find meaningfully significant and different effects of statues and placebo videos on passing exams after one month, and the estimated effect of placebo videos is very close to 0. For the six-month result, we are marginally able to reject the hypothesis of equal impacts of the statues and placebo videos with a p-value of 11% in the baseline specification in column 3 of panel (b).

As depicted in Table 4, we do not detect any effect on languages. In column 3 and 4, we find increments of 0.04 and 0.34 sd in the scores and exam passing in social sciences in statues group. These effects are significant at 5% only in panel (a). Nevertheless, we could not reject the null that the coefficients estimates for the impact of the statues and the placebo are not different from each other. As in panel (b), we do not find any significant effect after six months. Fearfulness towards Mathematics is prevalent in students aged 10-15 years (Soni and Kumari, 2017). Mathematics in particular, is also imperative for excelling in academics in that age. We propose that the impression of the role models made students focus courageously on important yet difficult tasks and perform better specifically on

Mathematics (and computer). These were messages communicated through the intervention videos. These in turn also validate the presence of a role modelling channel.

3.5.2 Heterogeneous treatment effects on boys and girls

In order to parse out differential treatment effects between boys and girls, we interact a variable containing that information with that of the treatment status. We find a strong and interesting association between gender and the extent of these effects. Boys and girls, in our sample, were very differently impacted by the intervention. We find that our intervention disproportionately impacted only the boys in the statues group. Table 4 shows a negative but insignificant effect on memory task for females in either of the statues or placebo group. While the statues group recorded significant improvements, they were completely driven by boys in that group. For the memory task, the overall effect of statues videos for male students is positive and significant 0.18 sd and its effect for female students is no longer significant (insignificant -0.02).

Likewise, the improvements in passing in the statues group were driven by boys. In Table 3, we find a strong and significant effect of 0.22 sd on passing for Math and Computer after a month. The effect amplified to 0.24 sd after six months. Column 2 and 3 of Table 5's first row show that boys recorded a higher and stronger improvement in passing the exam at 0.32 sd and 0.25 sd after one and six months. Although the effect after six months is weakly significant, the effect after a month is significant at 5%. While the overall effects on girls are not significant. We find a particularly strong effect on the performance of statues group students in STEM subjects[16], mostly driven by boys[17].

[16] A decomposition of these findings for different subjects are presented in Appendix TableA3.2. In addition of Mathematics and Computer, we also find a weakly significant 0.15 sd effect on English scores for both the placebo and placebo groups. There is no effect on the scores in Hindi although the statues and placebo videos were delivered in Hindi.

[17] A further decomposition of the HTE based on the subjects is presented in Appendix TableA3.4. As expected, for mathematics and computer the effect is run solely by boys in the statues group. Heterogeneity based on gender in English scores follows suit. Both the statues and placebo groups

At the bottom of Table 5, I also present the full effect of statues and placebo videos for each specification, respectively. Specifically, the full effect of statues videos refers to coefficient (A) + coefficient (B) * share of females in the sample. The results show that the effects of statues videos on memory task and exam outcomes of Math and Computer are statistically significant, while the effects of placebo videos are statistically insignificant. These results are consistent with those from the baseline specifications. As before, I show the p-values of the test for equality of the full effect of statues and placebo videos. Again, we cannot reject that statues and placebo videos have the same effect on the memory task immediately after the intervention, while we are (marginally) able to reject that they have equivalent effect on exam performance one month (six month) after the intervention.

A number of factors could explain the heterogeneous treatment effects. First, much in lines with the literature, females correspond only to female role models (Meier et al., 2020; Stout et al., 2011; Lockwood, 2006), unlike men who respond to role models regardless of their sex (Kipchumba et al., 2021). A dearth of female role models in our intervention may become a source for disparity in the form of inspiration that fuels effort and performance. Perhaps being predominantly surrounded by male role models signals to girls that conformity in the society may even mean not behaving like the other boys or sharing the same aspirations.

Second, the controversiality of our chosen role models may lead to the heterogeneous treatment effects by gender. Studies of brain activation during processing have documented greater activation in women than men in the left amygdala when processing negative emotional stimuli (Stevens and Hamann, 2012; Williams et al., 2005). In particular, Soroka (2014) has shown that negative trait ratings have a stronger effect on women's assessments of political leaders than men's. Soroka et al. (2016) find that women are more attentive than men to negatively framed political stories. In terms of the controversiality of the chosen role models in this study, Indira Gandhi may be one of the most salient ones. For instance, Prime Minister Indira Gandhi declared a state of emergency across India in 1975. During the

record a higher improvement in the scores of boys. There is no effect on the scores in Hindi.

Emergency, many of her political opponents were imprisoned and the press was censored. It has been widely seen as the darkest hour of India's democracy (Anderson and Clibbens, 2018; Indianexpress, 2015; Morris-Jones, 1975). As a result, girls participating in our experiment may be sensitive to the negative evaluations of Indira Gandhi, be reluctant to view her as a role model, and not be motivated by our intervention.

Third, the characteristics of the messages conveyed in the videos may explain the results. For example, the messages include many political elements, as most of the chosen role models are well-known because of their political influence. Considering that females are generally less interested in politics than men due to cultural, institutional, and family reasons, female students may simply be less inspired by the political messages in the videos than male students (Cowper-Coles, 2020). In addition, a small literature has detected gender differences in processing and evaluating information recorded with male and female voices (Linek et al., 2010; Kang et al., 2019). The messages for three male role models in our videos were conveyed by a male speaker. Therefore, female students may prefer female speakers and be less likely to be motivated by the videos than male students. Put together, there could be different reasons independent of the persons used to illustrate the messages that explain the heterogeneous treatment effects.

Lastly, the way to deliver the message may play a role in explaining the gender difference in the impact of our intervention. A branch of literature finds that there is a persistent patter of difference in the attitude, ability, and use of remote devices that favours males and boys in the learning context (Whitley Jr, 1997; Kay, 2008, 2009). Our intervention consisted of videos, which may result in a gender disparity by influencing boys only.

3.5.3 Robustness checks

A one-off treatment at an individual level preempts clustering by design. We still estimate equation 1 without and with controls (X_{it}). Apart from controlling baseline characteristics in all the estimates of treatment effects, we perform three robustness checks.

Firstly, we check if the attrition rate is associated with the outcome variables across the experiment groups. In Appendix TableA3.6 we find an absence of a significant relationship for the two groups. This ensures that the treatment did not relate or affect the rate at which students participated in the endline. Secondly, we also run a set of multiple hypothesis tests, following the method developed by Benjamini et al. (2006). Appendix TableA3.7 shows that the overall results are robust to multiple hypothesis testing. Particularly, the math and computer results remain significant. Thirdly, there is wide variation in the time used to complete the endline survey among participants. To take account of potential compliance issue, we have controlled for the time used to complete the endline survey in the main analysis. As a further robustness check, we examine whether the main results are sensitive to the exclusions of a small number of outliers regarding survey duration. Appendix TableA3.8 suggests that the magnitude and significance of the estimates is nearly identical through the varying specifications.

3.6 Discussion

3.6.1 Relevance of the role models

The famous figures are role models for students in many perspectives such as their personalities, experience, and actions. First, Mahatma Gandhi is known as the national father of India because of the role he played for freedom of the country. It was a very long and difficult journey of seeking his dream. He had to face many problems along the way. He spent over two thousand days in jail and suffered from several health problems during his journey (Jegen and Deats, 2005). Nevertheless, he never gave up, used positive thinking to overcome many difficulties and reduce stress (Gupta et al., 2019). His personality traits such as optimism and willpower may relate to the needs of students in India. National surveys have shown that a vast majority of primary and secondary students are suffering from math phobia in India (Times of India, 2014; Soni and Kumari, 2017; India Today, 2021). Students may learn from him and realize the strength of willpower. He may therefore act as a role model.

Second, Dr. Bhimrao Ramji Ambedkar was born an untouchable, and then became a famous figure in the history of India (Panyamanee, 2022). When he was in school, Ambedkar and other untouchable children were segregated and given no attention or assistance from the teachers. Even they were not allowed to sit inside the class (Zelliot, 2008; Sampathkumar, 2015). However, He studied hard to gain knowledge and respect. After he graduated from Bombay University, He received a scholarship to go to Columbia University and the London School of Economics (Zelliot, 2008). His life is a good example to illustrate how education could make a big difference in life. His experience may be relatable to disadvantaged students in India, and may update their beliefs about returns to getting education. Therefore, students may view him as a role model in life.

Third, Indira Gandhi is the first, and to this day, the only female Prime Minister of India (Malhotra, 2014). Gender inequality is still rife in many perspectives such as schooling and labor force participation in India (Dhar et al., 2022). Indira became an incredibly powerful force, even as a woman in a society dominated by patriarchal ideologies. Her career experiences may lead students to rethink their social position and potentially break stereotypes regarding gender roles (Beaman et al., 2012). Thus, she may be a role model as a counter example of traditional women.

Fourth, Arjuna is one of the chief protagonists of the Mahabharata. From a very young age, he started to show exemplary qualities of a student such as passion and commitment to learning. He practiced at night and showed activeness in his learning throughout. Hard work led him one day to all the strength and capability to become a great warrior, world class archer (Kelkar et al., 2022). Learning from Arjuna's experience, students may realise the importance of focus and dedication for them to achieve their goals and view him as a role mode.

Nevertheless, there were actually some aspects of their characteristics may not serve as the role models. For instance, an Emergency was imposed in India by the Prime Minister Indira Gandhi in 1975 that was very controversial, as it could suspend constitutional rights and restrict the democracy (Prakash, 2019). In addition, born into a prominent family, Indira was exposed to politics at an early age and led a highly politicized life, which may contribute to her success in politics (Green, 2013). However, such an experience may not be similar and attainable to many students. Students may think that she is just too far away from who they are. That is why she may not be a role model in this aspect. Also, most of the famous figures achieve their life goal in politics. However, some students may not be interested in political career. Therefore, they may not view them as role models. Overall, the famous figures, to a large extent, could be role models in terms of personality and experience, while in a number of aspects, they may not act as role models for students.

Relevance of the role models matters for the role-modelling effect. Ray (2006) develops the concept of an aspirations window, which is formed from an individual's cognitive world, her zone of "similar," "attainable" individuals. One individual draws her aspirations from the lives, achievements, or ideals of those who exist in her aspirations window and changes her behavior accordingly. Dalton et al. (2016) also highlight the importance of relevant role models for updating people's beliefs about their effort and aspirations. Students could reassess the feasible paths to success for them and change their beliefs about their ability to be successful based on the experience of role models (Riley, 2022). In the role models literature, there are mainly three dimensions of relatability.

First, people are motivated because of relevance in terms of physical characteristics such as sex and race. Many studies have shown that successful women can motivate other women (Beaman et al., 2012; Stout et al., 2011) and the role models from racial minorities matter for minorities (Aish et al., 2018). Second, people are influenced because of spatial relevance. There is some evidence that successful people cannot influence aspirations or behavior among the young after they emigrate from the inner city are no longer there to be observed (Wilson, 2012). Third, people are affected because of similar circumstances. Literature has shown that when people face threats such as cancer or marital breakup, they could be influenced by others in similar circumstances (Lockwood and Kunda, 1997). Taken together, relevance of the role models is important, as individuals use their peers (or near-peers) to form their aspirations and adjust their behavior.

3.6.2 Short run and long run effects

Statues could function as the repository of memory, commemorating role models and their inspiring stories, and then motivating students to emulate good deeds. Surprisingly, there is little literature in Economics on them and an absence of any experimental study on the relationship between statues and people's aspirations and behaviour. We do this precisely and find that such a relationship demands further exploration with very encouraging results. These results are not only of significance for broadening our understanding of statues and their role in society but also enhancing the literature on role models.

We think our statues intervention having an eventual role-modelling effect through the improved familiarity with and susceptibility to the role models. Statues and messages conveyed in the videos may help to invoke feelings of familiarity with and susceptibility to the role models.

First, drawing from corporate brand literature (Perera and Chaminda, 2013), familiarity refers to cumulative knowledge about the role models that is acquired through exposure to the role models. Familiarity with the role models is the precondition for individuals to spot the common characteristics with that of the role model and to be influenced by the role model. Literature on psychology about brand familiarity has shown that an automatic frequency counting mechanism exists in memory, and frequent exposure to the brand can influence consumers' preference and behaviors (Hasher and Zacks, 1984; Baker et al., 1986). Importantly, literature on the role models has emphasized the frequency of exposure to the role models and found that frequent contact with the role models can lead to stronger self-esteem and more career ambitions (Asgari et al., 2010). In our study, students in statues group becomes more familiar with the role models by attending the virtual tour to statues, as locally presented role models in the form of statues could have a memory stimulus on students.

Second, familiarity only may not enough to activate a role-modelling effect. Based on the literature (Laursen and Faur, 2022), susceptibility refers to the willingness for individuals to identify the common characteristics with that of the role model and further conform. It could be an important determinant of inducing the role-modelling effect. In particular, the studies on social networks have shown that individuals' susceptibility to social influencers is one key factor for effective social influence and behavior changes (Watts and Dodds, 2007; Aral and Walker, 2012; Stöckli and Hofer, 2020). Moreover, literature on heritage has suggested that positive emotions are associated with destination experience at heritage sites such as museums (Tung and Ritchie, 2011; Prayag et al., 2013; Isaac and Budryte-Ausiejiene, 2015). Likewise, the virtual tour to statues may induce positive emotions, and therefore make students become more susceptible to the role models. That is why we introduce susceptibility in our context. However, we did not measure the susceptibility in the data, and it is only a hypothesis.

A memory stimulus and increase in career aspirations that are pro-socially driven for public service suggest the presence of vicarious effects. These rely on two pillars. Firstly, we invoke familiarity by presenting carefully tailored stories that these role models lived (our narratives)[18]. Secondly, by developing susceptibility (only for the statues group) for these statues in the near future using associative memory (Kahneman, 2011). However, our data analysis is not able to show the associative memory of students relating to statues. Therefore, it is only a hypothesis. These are presented in Figure 3.

The one-month effects include an improvement in academic performance (scores

[18] And in the statues group only, showcasing easily accessible statues (visual cues that are appealing) of these role models

and number of students passing the exam) on Mathematics and Computer, and Social Sciences. These improvements are markedly significant only for the statues group[19]. The longer run effects mimic these to some extent for the passing rate in Math and Computer. These results are strong, significant and special due to the fact that the intervention was merely a 6-minute video (or virtual tour for the statues group students).

3.6.3 Mechanism

Aesthetic experiences affect psyche and mood, and promote health and well-being by eliciting positive emotional outputs (Mastandrea et al., 2019). Statues are one such subtle aesthetic experience that if conditioned strategically can influence individual motivation and performance. We employ the concept of role-modelling as the best fit to explore such an influence.

We find that our intervention had an effect on student effort and academic performance. While the intervention tool in itself is a role-modelling video, we scrtunitize the constituting pieces of the intervention to explain the mechanisms. In that attempt we unearth two key components that may be fundamental for relatability to different role models. As part of the experimental design we manipulate this facet.

Any role-modeling intervention may depend if the sample/viewer's can (i) relate to the protagonists and (ii) the level of their goals and the one's attained by the protagonists. We delve deeper into the first constituent of relatability or relevance of the role-model. We suggest that it may rely on two key constituents, namely familiarity and susceptibility.

Familiarity suggests possessing knowledge about the role model. The amount of knowledge may vary depending upon priors and information exposure. This may help individuals assimilate information about themselves and identify the

[19]We do have some significant estimate for the placebo group too. These include the scores in English after a month and the improvements in career aspirations, much alike the statues group. This does not come as a surprise, as we do expect that placebo video to still have some -albeit weaker-role modelling effect. The rationale for such an expectation is summed up in Figure 3.

overlapping characteristics with that of the role model. Higher the overlap the more they may relate with the role model. The higher the familiarity, the more characteristics may they explore to examine this overlap (as an automatic process). Familiarity therefore, may offer a wider array of aspects of personality traits and life experiences to match between the role models and the subject based on their spectrum of such traits, worldviews and experiences.

Susceptibility, captures a likelihood of being influenced in light of new information or experiences. Within the framework of relatability, the more susceptible an individual is, the more likely are to try and identify the above-mentioned overlaps. Higher susceptibility may correspond to an increased chance of willingly identified overlaps resulting in higher levels of perceived relevance of the role model. Together, familiarity and susceptibility may determine the level at which any subject relates (voluntarily and involuntarily, knowingly and unknowingly, consciously and unconsciously[20]) to a role model.

While familiarity may offer to determine how relatable a role model will be by expanding the spectrum of traits and experiences, susceptibility may warrant the will to identify overlaps out of this set of traits, vignettes and life-experiences. In close conformity, these two may be the building blocks of any role-modelling effect along with the goals set and attained by these role models.

In this RCT, using statues we maneuver subtle manipulations in these fine-grain constituents within relatability. In the statues group, we show a virtual tour of four statues to which, the students in the sample might possess some amounts of familiarity, though we did not measure their initial familiarity in the data. This is because four role models are introduced in their textbooks and they may learn them in the curriculum (see Figure 4 to Figure 7 for students' textbooks about the role models). In the placebo group, we replace the statues with images of the same role models. This allows to take one element of familiarity away from the placebo group. Having attended the virtual tour, students in the statues group are more familiar with the statues and in turn the role models that they embody. Likewise, with the

[20]Resulting from the two systems that are automatic and deliberative at work (Kahneman, 2011).

same manipulation, we heighten the susceptibility of the statues group students to these statues for future. This feature is missing in the placebo or control groups. Lastly, using a meticulously scripted narrative that is played in the background of both the groups' video, we balance the goals-attained by these role models. Placebo and statues groups hear the same narrative where role models in both groups are shown to attain the same goals. By the virtue of this design, students in the statues group end up being more familiar and susceptible to the statues and the role models they represent. This translates into the treatment effects that we discussed in section 5.

Figure 3 displays the conceptual framework of this study and offers a roadmap for the channels. Immediately after the intervention, heightened relatability (stemming from higher familiarity and susceptibility) results in higher aspiration and effort in a memory task. Dalton et al. (2016) show that aspirations and effort are co-determined and our findings echo this. These improvements later reflect in higher academic performance.

In total, there are five hypotheses in Figure 3. First, the statues group feels more familiar with statues after the intervention. We test this hypothesis by exploring the effect of our intervention on a memory test. We show that our statues intervention has an instant memory stimulus among students. However, our data analysis cannot test its impact in a relatively long term. Second, both groups are equally familiar to the role models accessible as a part of their curriculum. This is an untestable hypothesis, as we did not measure students' initial knowledge about the role models in the data.

Third, treated students become more susceptible to the role models that are readily accessible and available in the form of statues. This is an untestable hypothesis, as we did not capture students' willingness to identify the overlapping characteristics with the role models or their emotional feelings for them in the data. Fourth, identical narrative exposes the students across both groups to the same goals. We test this hypothesis by investigating the impact of our intervention on students' career aspirations for public service. We find that both statues and placebo videos have statistically significant and positive influence on their career aspirations for public service. Lastly, aspirations and effort are co-determined. We do not directly test this hypothesis. Nevertheless, this connection has been well-documented in the literature (Dalton et al., 2016; Fruttero et al., 2021).

Up until January 2022, primary school students in Jaipur had online classes. This restricted their contact with the statues. However, this changed when schools were re-opened in February. With the easing of lockdown restrictions students regained frequent access to the statues (that only the statues group students had visited virtually) among every other place of visit. A month before the examinations, only the statues students were more familiar to these statues. We justify our results on the premise that the intervention had strong effects immediately after the intervention that lasted after 6 months due to increased familiarity.

Many role-modelling interventions fall short of having a desired effect. Evidence and explanation for this is mixed. Appreciation for context, culture and knowledge of the two fundamentals of relatability and goal-setting have enabled researchers and policy makers to better equip themselves. Complemented with experimental evidence we further this understanding. Surely, more research in this direction will be largeley beneficial as one of the key advantage of such interventions is their cost-effectiveness. To highlight this, to produce the intervention we spent a little over 800 GBP[21].

3.7 Conclusion

Statues are one of the strongest and loudest expressions of a society/nation's power and pride and paradoxically, an immediate source of their vulnerability and incongruity. A peripheral view of the costs and benefits associated with erection of a statue is usually limited to the costs of material and maintenance, and the employment or revenue from tourism they may generate.

[21]The statues were free. The same applied to the voice-over artists. We are extremely grateful to the Rajasthan Police and the Traffic Police Department, Government of Rajasthan for their support with the production of the intervention.

A deeper view may also involve their existence benefit for the local community via pride and utility gains (in the form of their aesthetics translating into general well-being). Their option value involves the gains to those who do not directly consume their benefits but may wish to preserve these to reap potential future benefits (Benhamou, 2020). The opportunity cost of having something less aesthetically appealing to replace them can also be taken into consideration. Nevertheless, beyond these, we introduce two distinct constituents to this cost-benefit equation for policy makers.

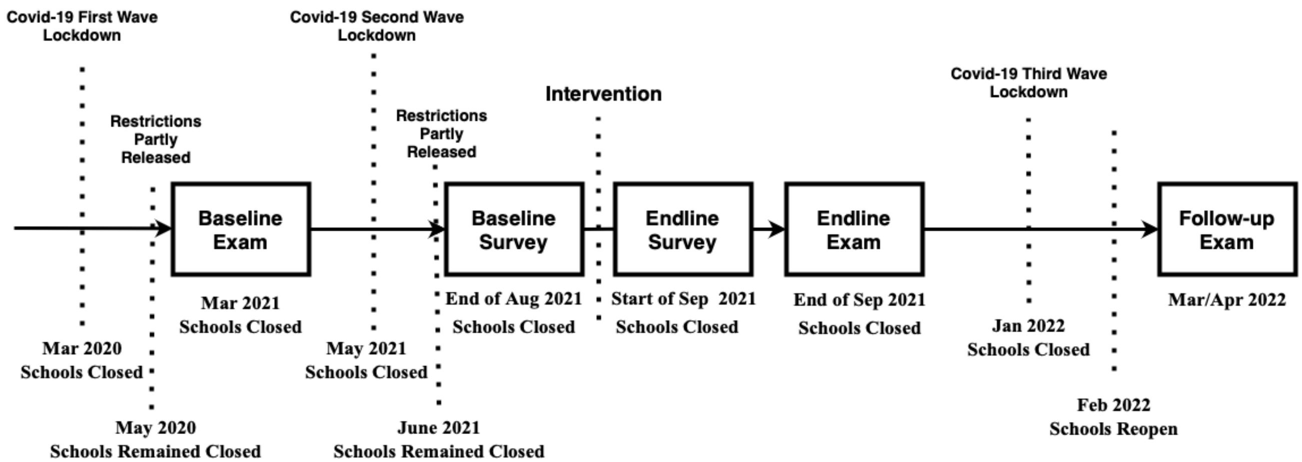
First, an inspirational role-modelling effect contributes to early child development and improved human capital accumulation. Such motivational effects can be manifold and we barely scratch the surface by limiting our attention to a sample of primary school children. It would be useful to assess similar effects across other dimensions and in different settings.

Second, we show that role model intervention could give rise to gender disparities. We find that our intervention causes a memory stimulus and improvement in effort only for boys. There could be multiple explanations for this finding such as lack of female role models, existence of controversiality of the chosen role models, and/or absence of interests in politics for females.

The limitation of this work is that no further surveys were conducted with the students, preventing a deeper understanding of the change in the additional memory stimulus in a relatively long horizon after the intervention. Particularly, we cannot reject that statues and placebo interventions have the same effect on the memory stimulus immediately after the intervention, though the placebo intervention on its own does not have statistically significant effect and its coefficient is much smaller than that of the statues in magnitude.

Statues are important commemorative tools that bear aesthetic beauty, historical information and implications on socioeconomic development of the society. They should be harnessed carefully to overcome stereotypes, disparities and foster harmony and development. Statues can help inspire and if installed and maintained sensitively, bridge gaps within our society. Our study opens avenues for future research and broadens our understanding of historic monuments and role models.

Fig. 1: Timeline



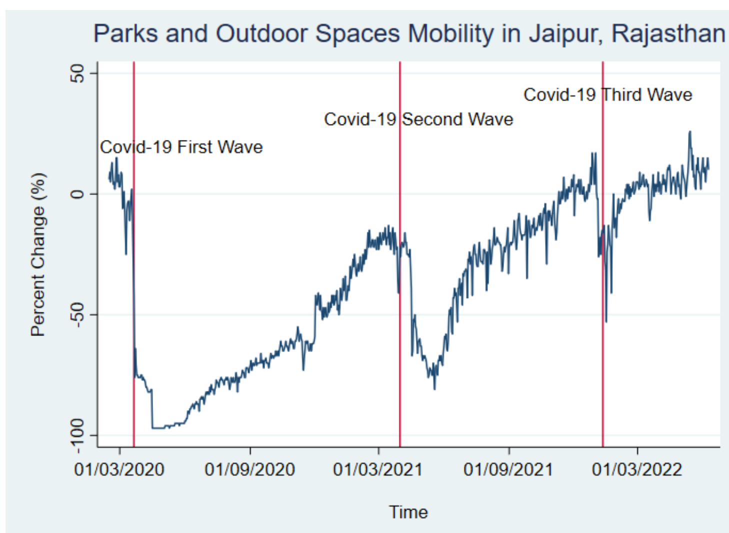


Fig. 2: Experiment Assignment

Statues	**Placebo**	**Control**
Statues -virtual tour	Motivational video	No Intervention
Narrative	Narrative	

Table 1: Balance Table

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	Statues		Placebo		Control		
	Obs	Mean	Obs	Mean	Obs	Mean	p-value
Female	538	0.42	513	0.41	520	0.42	0.91
Age	538	9.43	513	9.48	520	9.46	0.61
Class 4	538	0.54	513	0.51	520	0.49	0.28
Duration	523	1.15	523	1.44	524	1.15	0.74
Device (Computer)	523	0.23	523	0.23	524	0.23	0.99
Memory Score	491	5.26	482	5.04	488	5.09	0.56
Public Career Aspiration	393	0.74	395	0.76	389	0.75	0.73
Hope	517	24.63	517	24.82	520	25.24	0.19
Aspiration	511	21.39	510	21.28	517	21.04	0.30
Math Score	538	0.85	513	0.85	520	0.84	0.65
Computer Score	396	0.84	375	0.85	388	0.84	0.61
English Score	537	0.82	513	0.83	520	0.81	0.29
Hindi Score	538	0.87	513	0.87	520	0.86	0.28
Social Science Score	458	0.84	428	0.85	446	0.84	0.52

Notes: This table examines whether baseline characteristics and main outcomes are balanced across treatment, placebo, and control groups. Column 7 shows an F-test of equality of the means across the three groups for each characteristic and outcome.

Fig. 3: Conceptual Framework

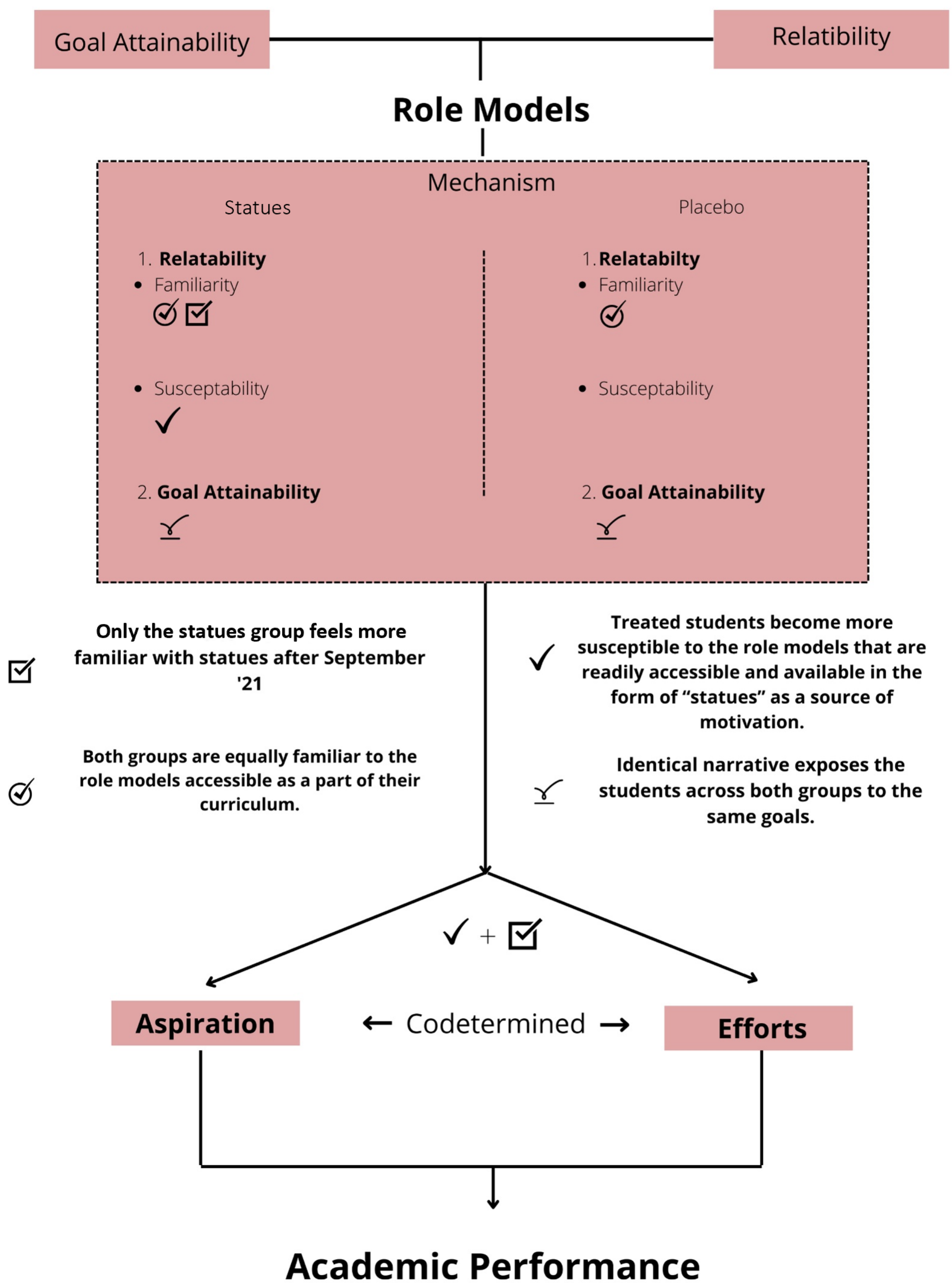


Table 2: Estimates of the Effect of Statues (Short Term) on memory stimulus and aspirations

	Memory Test		Public Service Aspiration	
	(1)	(2)	(3)	(4)
Statues	0.139**	0.119*	0.111**	0.143**
	(0.062)	(0.066)	(0.056)	(0.061)
Placebo	0.075	0.052	0.106*	0.137**
	(0.065)	(0.067)	(0.055)	(0.063)
P-value: Statues=Placebo	0.31	0.29	0.92	0.91
Controls	NO	YES	NO	YES
Observations	981	981	790	790

Notes: This table reports estimates of the effect of treatment on students' memory test scores and public career aspiration immediately after the intervention. All outcomes are standardised. Controls include gender, age, class, device used to watch the videos, and the time used to complete the endline survey. Heteroskedasticity robust standard errors are reported below point estimates. Significance at the 10%, 5%, and 1% levels is indicated by *, **, and ***, respectively.

Table 3: Estimates of the Effect of Statues on Mean Score of Math and Computer (One/Six-Month)

	Score		Pass	
	(1)	(2)	(3)	(4)
Panel A: One Month				
Statues	0.207***	0.211***	0.218*	0.225**
	(0.080)	(0.082)	(0.127)	(0.114)
Placebo	0.099	0.114	-0.004	0.013
	(0.089)	(0.090)	(0.160)	(0.158)
P-value: Statues=Placebo	0.18	0.22	0.08	0.09
Controls	NO	YES	NO	YES
Observations	1,026	1,026	1,026	1,026
Panel B: Six Month				
Statues	0.065	0.091	0.225**	0.240**
	(0.080)	(0.085)	(0.099)	(0.115)
Placebo	0.064	0.096	0.097	0.115
	(0.084)	(0.087)	(0.133)	(0.147)
P-value: Statues=Placebo	0.98	0.94	0.11	0.12
Controls	NO	YES	NO	YES
Observations	1,001	1,001	1,001	1,001

Notes: This table reports estimates of the effect of treatment on students' math and computer performance. Panel A reports the effect one month after the intervention, while panel B reports the effect six months after the intervention. The dependent variable is the mean score of math and computer exams in Columns 1 and 2, while the dependent variable is a dummy for whether a student passed the exams in columns 3 and 4. All outcomes are standardised. Controls include gender, age, class, device used to watch the videos, and the time used to complete the endline survey. Heteroskedasticity robust standard errors are reported below point estimates. Significance at the 10%, 5%, and 1% levels is indicated by *, **, and ***. respectively.

Table 4: Estimates of the Effect of Statues on Language and Social Science (One/Six-Month)

	Language		Social Science	
	Score	Pass	Score	Pass
	(1)	(2)	(3)	(4)
Panel A: One Month				
Statues	0.108	0.191	0.039**	0.336**
	(0.083)	(0.130)	(0.016)	(0.161)
Placebo	0.108	0.089	0.025	0.217
	(0.088)	(0.147)	(0.017)	(0.177)
P-value: Statues=Placebo	0.99	0.43	0.35	0.39
Controls	YES	YES	YES	YES
Observations	1051	1051	903	903
Panel B: Six Month				
Statues	0.032	-0.071	0.073	-0.173
	(0.077)	(0.095)	(0.087)	(0.106)
Placebo	0.008	-0.180	0.020	-0.186
	(0.086)	(0.127)	(0.093)	(0.125)
P-value: Statues=Placebo	0.78	0.38	0.56	0.92
Controls	YES	YES	YES	YES
Observations	1019	1019	879	879

Notes: This table reports estimates of the effect of treatment on the mean score of students' language exams including English and Hindi, and social science exam score. Panel A reports the effect one month after the intervention, while panel B reports the effect six months after the intervention. The dependent variable is exam score in Columns 1 and 3, while the dependent variable is a dummy for whether a student passed the exam in columns 2 and 4. All outcomes are standardised. Controls include gender, age, class, device used to watch the videos, and the time used to complete the endline survey. Heteroskedasticity robust standard errors are reported below point estimates. Significance at the 10%, 5%, and 1% levels is indicated by *, **, and ***.

Table 5: Heterogeneity in Treatment Effect by Gender

	Memory Test	Math/Computer Passing	
	Immediate	One-Month	Six-Months
	(1)	(2)	(3)
(A) Statues	0.216**	0.317**	0.245*
	(0.085)	(0.146)	(0.147)
(B) Statues * Female	-0.234*	-0.211	-0.013
	(0.124)	(0.260)	(0.199)
(C) Placebo	0.144	0.236	0.107
	(0.088)	(0.157)	(0.200)
(D) Placebo * Female	-0.220*	-0.532	0.019
	(0.131)	(0.347)	(0.261)
Female	0.112	0.015	0.027
	(0.090)	(0.225)	(0.197)
Controls	YES	YES	YES
Observations	981	1026	1001
Overall Treatment Effect			
Female in Statues (A) + (B)	-0.018	0.106	0.232
	(0.085)	(0.201)	(0.158)
Female in Placebo (C) + (D)	-0.076	-0.296	0.127
	(0.100)	(0.309)	(0.189)
Full Effect			
Statues	0.118*	0.229**	0.240**
	(0.065)	(0.114)	(0.116)
Placebo	0.053	0.017	0.115
	(0.067)	(0.156)	(0.148)
P-value: Statues=Placebo	0.31	0.09	0.12

Notes: This table reports estimates of the effect of treatment on students' memory test scores and math/computer passing by student gender. All outcomes are standardised. Controls include age, class, device used to watch the videos, and the time used to complete the endline survey. The bottom panel shows the overall treatment effect for each sub-group. Heteroskedasticity robust standard errors are reported below point estimates. Significance at the 10%, 5%, and 1% levels is indicated by *, **, and ***, respectively.

Fig. 4: Mahatma Gandhi

Notes: This photo is from the school teacher showing that Mahatma Gandhi is introduced in students' textbook.

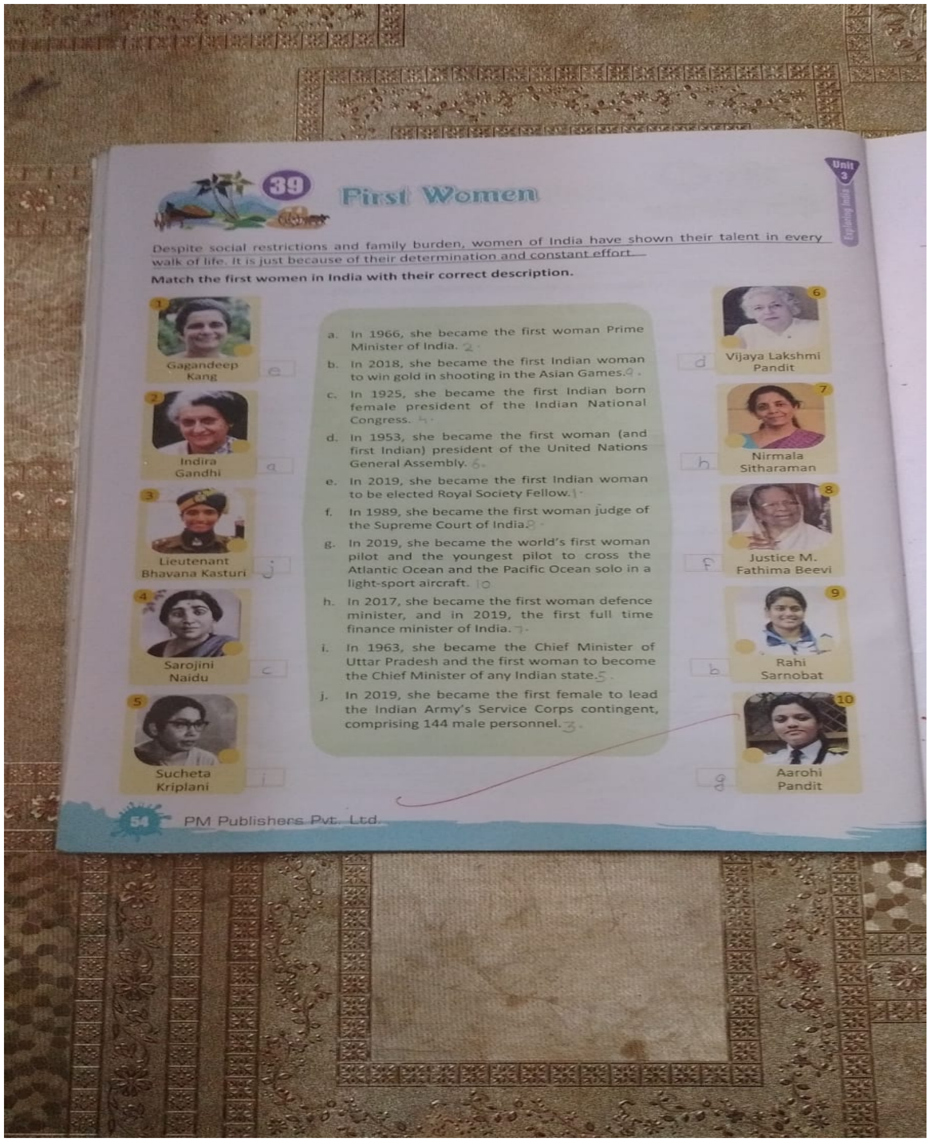
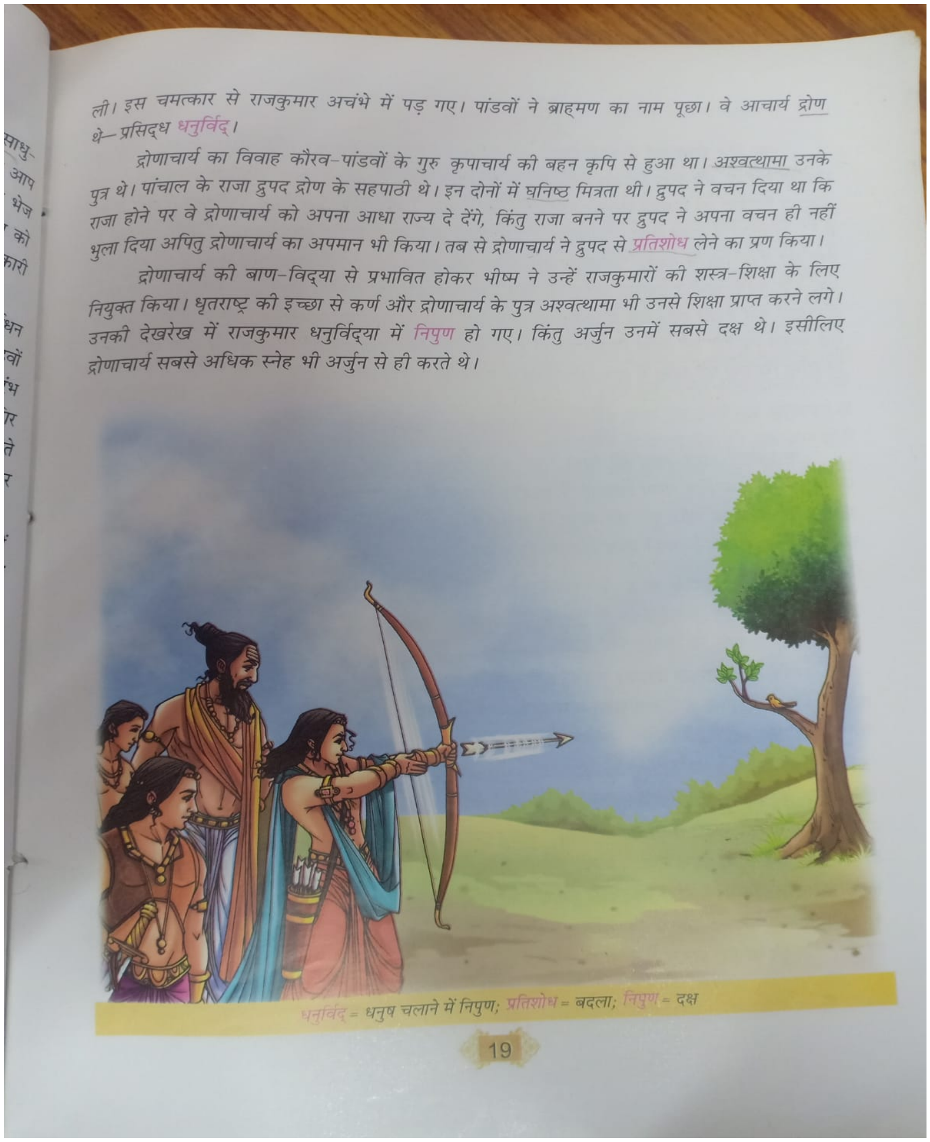


Fig. 5: Indira Gandhi

Notes: This photo is from the school teacher showing that Indira Gandhi is introduced in students' textbook.



ली। इस चमत्कार से राजकुमार अचंभे में पड़ गए। पांडवों ने ब्राह्मण का नाम पूछा। वे आचार्य द्रोण थे– प्रसिद्ध धनुर्विद्।

द्रोणाचार्य का विवाह कौरव-पांडवों के गुरु कृपाचार्य की बहन कृपि से हुआ था। अश्वत्थामा उनके पुत्र थे। पांचाल के राजा द्रुपद द्रोण के सहपाठी थे। इन दोनों में घनिष्ठ मित्रता थी। द्रुपद ने वचन दिया था कि राजा होने पर वे द्रोणाचार्य को अपना आधा राज्य दे देंगे, किंतु राजा बनने पर द्रुपद ने अपना वचन ही नहीं भुला दिया अपितु द्रोणाचार्य का अपमान भी किया। तब से द्रोणाचार्य ने द्रुपद से प्रतिशोध लेने का प्रण किया।

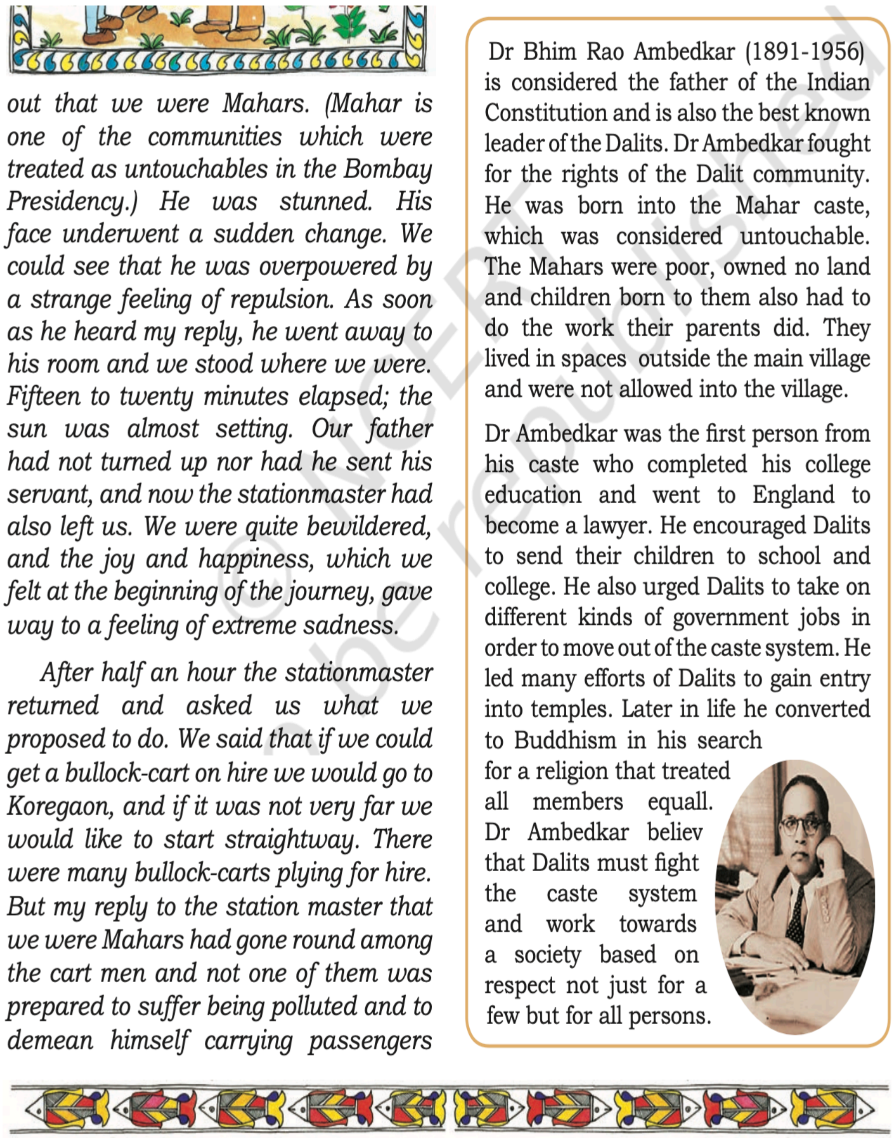
द्रोणाचार्य की बाण-विद्या से प्रभावित होकर भीष्म ने उन्हें राजकुमारों की शस्त्र-शिक्षा के लिए नियुक्त किया। धृतराष्ट्र की इच्छा से कर्ण और द्रोणाचार्य के पुत्र अश्वत्थामा भी उनसे शिक्षा प्राप्त करने लगे। उनकी देखरेख में राजकुमार धनुर्विद्या में निपुण हो गए। किंतु अर्जुन उनमें सबसे दक्ष थे। इसीलिए द्रोणाचार्य सबसे अधिक स्नेह भी अर्जुन से ही करते थे।

धनुर्विद् = धनुष चलाने में निपुण; प्रतिशोध = बदला; निपुण = दक्ष

19

Fig. 6: Arjuna

Notes: This photo is from the school teacher showing that Arjuna is introduced in students' textbook.



out that we were Mahars. (Mahar is one of the communities which were treated as untouchables in the Bombay Presidency.) He was stunned. His face underwent a sudden change. We could see that he was overpowered by a strange feeling of repulsion. As soon as he heard my reply, he went away to his room and we stood where we were. Fifteen to twenty minutes elapsed; the sun was almost setting. Our father had not turned up nor had he sent his servant, and now the stationmaster had also left us. We were quite bewildered, and the joy and happiness, which we felt at the beginning of the journey, gave way to a feeling of extreme sadness.

After half an hour the stationmaster returned and asked us what we proposed to do. We said that if we could get a bullock-cart on hire we would go to Koregaon, and if it was not very far we would like to start straightway. There were many bullock-carts plying for hire. But my reply to the station master that we were Mahars had gone round among the cart men and not one of them was prepared to suffer being polluted and to demean himself carrying passengers

Dr Bhim Rao Ambedkar (1891-1956) is considered the father of the Indian Constitution and is also the best known leader of the Dalits. Dr Ambedkar fought for the rights of the Dalit community. He was born into the Mahar caste, which was considered untouchable. The Mahars were poor, owned no land and children born to them also had to do the work their parents did. They lived in spaces outside the main village and were not allowed into the village.

Dr Ambedkar was the first person from his caste who completed his college education and went to England to become a lawyer. He encouraged Dalits to send their children to school and college. He also urged Dalits to take on different kinds of government jobs in order to move out of the caste system. He led many efforts of Dalits to gain entry into temples. Later in life he converted to Buddhism in his search for a religion that treated all members equall. Dr Ambedkar believ that Dalits must fight the caste system and work towards a society based on respect not just for a few but for all persons.

Fig. 7: B.R. Ambedkar

Notes: This photo is from the school teacher showing that Dr. B.R. Ambedkar is introduced in students' textbook.

Table A3.1: Summary Statistics

	Label	Observations	Mean	SD	Min	Max
Female	0 - male, 1 - female	1,571	0.42	0.49	0	1
Age	# years	1,571	9.46	0.79	5	12
Class 4	0 - class 5, 1 - class 4	1,571	0.51	0.50	0	1
Duration	# hours	1,570	1.25	7.24	0.07	169.86
Device (Computer)	0 - others, 1 - computer	1,570	0.23	0.42	0	1
Memory Score	0 - 10	1,461	5.13	3.29	0	10
Public Career Aspiration	0 - other, 1 - public career	1,177	0.75	0.44	0	1
Hope	0 - 36	1,554	24.90	5.50	6	36
Aspiration	0 - 24	1,538	21.24	3.75	3	24
Math Score	school exam	1,571	0.84	0.15	0	1
Computer Score	school exam	1,159	0.85	0.16	0	1
English Score	school exam	1,570	0.82	0.16	0	1
Hindi Score	school exam	1,571	0.86	0.14	0	1
Social Science Score	school exam	1,332	0.84	0.17	0	1

Notes: This table displays observations, means, standard deviations as well as the minimum and maximum values for variables of interest.

Table A3.2: Estimates of the Effect of Statues by Subject (One/Six-Month)

	Math	Computer	English	Hindi
	(1)	(2)	(3)	(4)
Panel A: One Month				
Statues	0.180*	0.223**	0.149*	0.069
	(0.092)	(0.097)	(0.080)	(0.087)
Placebo	0.083	0.116	0.154*	0.069
	(0.103)	(0.107)	(0.085)	(0.093)
P-value: Statues=Placebo	0.30	0.27	0.94	0.99
Controls	YES	YES	YES	YES
Observations	1053	796	1051	1053
Panel B: Six Month				
Statues	-0.012	0.045	0.079	-0.007
	(0.088)	(0.093)	(0.073)	(0.084)
treat2	0.031	-0.026	0.040	-0.003
	(0.095)	(0.099)	(0.081)	(0.093)
P-value: Statues=Placebo	0.64	0.45	0.63	0.96
Controls	YES	YES	YES	YES
Observations	1020	771	1019	1021

Notes: This table reports estimates of the effect of treatment on exam score by subject. Panel A reports the effect one month after the intervention, while panel B reports the effect six months after the intervention. All outcomes are standardised. Controls include gender, age, class, device used to watch the videos, and the time used to complete the endline survey. Heteroskedasticity robust standard errors are reported below point estimates. Significance at the 10%, 5%, and 1% levels is indicated by *, **, and ***.

Table A3.3: Estimates of the Effect of Statues (Short Term)

	Hope	Aspiration	Happy	Life Satisfaction
	(1)	(2)	(3)	(4)
Statues	0.099	-0.082	-0.031	-0.100
	(0.064)	(0.081)	(0.075)	(0.075)
Placebo	0.057	0.057	-0.020	-0.095
	(0.062)	(0.074)	(0.065)	(0.072)
P-value: Statues=Placebo	0.47	0.05	0.87	0.94
Controls	YES	YES	YES	YES
Observations	1048	1041	1039	1036

Notes: This table reports estimates of the effect of treatment on students' psychological outcomes immediately after the intervention. All outcomes are standardised. Controls include gender, age, class, device used to watch the videos, and the time used to complete the endline survey. Heteroskedasticity robust standard errors are reported below point estimates. Significance at the 10%, 5%, and 1% levels is indicated by *, **, and ***, respectively.

Table A3.4: Gender Heterogeneity in Treatment Effect on Exam Score by Subject

	One-Month			
	Math	Computer	English	Hindi
	(1)	(2)	(3)	(4)
(A) Statues	0.306***	0.330**	0.177*	0.138
	(0.118)	(0.135)	(0.106)	(0.118)
(B) Statues * Female	-0.301	-0.165	-0.065	-0.165
	(0.184)	(0.191)	(0.159)	(0.171)
(C) Placebo	0.033	0.217	0.194*	0.045
	(0.134)	(0.139)	(0.109)	(0.119)
(D) Placebo * Female	0.118	-0.143	-0.094	0.055
	(0.207)	(0.221)	(0.170)	(0.184)
Female	0.006	0.132	0.108	0.133
	(0.143)	(0.148)	(0.126)	(0.134)
Controls	YES	YES	YES	YES
Observations	1053	873	1051	1053
Overall Treatment Effect				
Female in Statues (A) + (B)	0.005	0.164	0.112	-0.027
	(0.143)	(0.139)	(0.121)	(0.126)
Female in Placebo (C) + (D)	0.151	0.074	0.100	0.101
	(0.160)	(0.173)	(0.132)	(0.143)
Full Effect				
Statues	0.179*	0.260***	0.149*	0.069
	(0.092)	(0.098)	(0.080)	(0.087)
Placebo	0.082	0.158	0.155*	0.068
	(0.103)	(0.109)	(0.085)	(0.093)
P-value: Statues=Placebo	0.29	0.29	0.93	0.99

Notes: This table reports estimates of the effect of treatment on students' exam scores by gender one month after the intervention. All outcomes are standardised. Controls include age, class, device used to watch the videos, and the time used to complete the endline survey. The bottom panel shows the overall treatment effect for each sub-group. Heteroskedasticity robust standard errors are reported below point estimates. Significance at the 10%, 5%, and 1% levels is indicated by *, **, and ***, respectively.

Table A3.5: Attrition Table

	(1)	(2)	(3)	
	Statues Group	Images Group	Control Group	
	Obs	Mean	Obs	*p*-value
Baseline	538	513	520	
Attrition Rate	0.27	0.26	0.28	0.70

Notes: This table examines the attrition rate by groups. Column 4 shows an F-test of equality of the means across the three groups for attrition.

Table A3.6: Attrition Regression

	(1)
	Attrition
Statues	-0.013
	(0.028)
Placebo	-0.023
	(0.028)
Observations	1,571

Notes: Linear regression of treatment indicators on a variable equal to one if the student was not surveyed at endline. Robust standard errors in parentheses.

Table A3.7: Multiple Hypothesis Tests for Estimates of the Effect of Statues (One/Six-Month)

	Exam Passing		
	Math/Computer	Language	Social Science
	(1)	(2)	(3)
Panel A: One Month			
Statues	0.225**	0.191	0.336**
	(0.114)	(0.130)	(0.161)
	[0.099]*	[0.285]	[0.074]*
Placebo	0.013	0.089	0.217
	(0.158)	(0.147)	(0.177)
	[1.000]	[1.000]	[0.439]
P-value: Statues=Placebo	0.09	0.43	0.39
Controls	YES	YES	YES
Observations	1,026	1,051	903
Panel B: Six Month			
Statues	0.240**	-0.071	-0.173
	(0.115)	(0.095)	(0.106)
	[0.076]*	[0.915]	[0.205]
Placebo	0.115	-0.180	-0.186
	(0.147)	(0.127)	(0.125)
	[0.866]	[0.312]	[0.279]
P-value: Statues=Placebo	0.12	0.38	0.92
Controls	YES	YES	YES
Observations	1,001	1,019	879

Notes: This table reports estimates of the effect of treatment on exam passing with q-value. All outcomes are standardised. Controls include gender, age, class, device used to watch the videos, and the time used to complete the endline survey. Heteroskedasticity robust standard errors are reported below point estimates. Significance at the 10%, 5%, and 1% levels is indicated by *, **, and ***, respectively. Sharpened q-value in square brackets.

Table A3.8: Sensitivity Test for Excluding Outliers of Survey Duration (One/Six-Month)

	Math/Computer Exam Passing Duration		
	< 1.5 Hours	< 2 Hours	< 4 Hours
	(1)	(2)	(3)
Panel A: One Month			
Statues	0.266**	0.249**	0.233**
	(0.120)	(0.119)	(0.118)
Placebo	0.031	0.030	0.013
	(0.166)	(0.165)	(0.163)
P-value: Statues=Placebo	0.08	0.09	0.09
Controls	YES	YES	YES
Observations	1,011	1,022	1,026
Panel B: Six Month			
Statues	0.270**	0.262**	0.249**
	(0.130)	(0.127)	(0.120)
Placebo	0.133	0.133	0.120
	(0.160)	(0.159)	(0.152)
P-value: Statues=Placebo	0.11	0.12	0.12
Controls	YES	YES	YES
Observations	986	997	1,001

Notes: This table reports estimates of the effect of treatment on exam passing by the time used to complete the endline survey. All outcomes are standardised. Controls include gender, age, class, device used to watch the videos, and the time used to complete the endline survey. Heteroskedasticity robust standard errors are reported below point estimates. Significance at the 10%, 5%, and 1% levels is indicated by *, **, and ***, respectively.

3.7.1 Intervention script

M.K. Gandhi

Mahatma Gandhi is known as the father of our nation. He led India to freedom against the British. His legacy, however, has a humble beginning.

One night, he was boarding a train in South Africa. He had a ticket for a first class compartment. When he went to his seat, he was told that the compartment was for whites only. He was thrown out.

He spent the whole night in the cold, thinking about how much discrimination he would face in the future. He thought: he can either accept reality - as it is - give up his career, and go back. Or he can stay, and fight for his rights and those of his people!

He spent 21 years in South Africa fighting against discrimination, and then went back to free India from the British. He led the Satyagraha Movement. In spite of all the challenges that came his way, his faith was unfettered. As he persevered, people joined him in his quest.

He had to face many problems along the way. He spent over two thousand days in jail, and suffered from several health problems during his journey.

The dream of India's freedom was the only thing in his mind. It was because of this determination and will power that we are living a free and happy life in our Independent India!

"The future depends on what you do today."

Arjuna

Arjuna - the son of Kunti and Pandu - was one of the most important characters of the Mahabharata. From a very young age, his passion and commitment to learning was clear.

Never missing an opportunity to learn, one evening, while eating in the dark, he realised that his hands reached his mouth effortlessly. He figured, if practice could help him eat in the dark, it could also help him aim in the dark. He started practicing archery immediately, and perfected his aim.

Later one day, Guru Dronacharya - the teacher of the Pandavas - asked them to aim and shoot at the eye of a wooden fish that he had tied to a tree. They had to shoot by looking at its reflection in the lake below. He then asked each student what they saw. Their answers ranged from the tree, water, leaves, to the sky. But Arjuna said: "The eye."

Arjuna could disregard everything else, and focus his mind and attention only on his target. Naturally, the shot was executed successfully.

Through focus and dedication, Arjuna emerged as one of the strongest and most skilled characters in the Mahabharata.

"When you want to achieve something, you must focus on it. Close out all other distractions, and concentrate only on your target." -Guru Dronacharya

Indira Gandhi

Indira Gandhi, the Iron Lady of India, was the first, and to this day, the only female Prime Minister of India. She brought prosperity to the general public and strength to the nation.

Her father, Dr. Jawaharlal Nehru, was often away, and her mother, frequently sick. She spent a lonely childhood. But her father used to write her letters full of wisdom and inspiration. In one letter he wrote-

"For the desire to hide anything means that you are afraid, and fear is a bad thing and unworthy of you. Be brave, and all the rest follows. We work in the sun and in the light."

These letters had a significant impact on her personality, bold and courageous.

Growing up, she didn't let gender stereotypes stop her. She grew up climbing trees, flying kites, and playing marbles with her cousins, most of whom were male.

Indira Gandhi stunned the whole world with her fearlessness, confidence, and strength by breaking the glass ceiling!

"Let's not forget, in India the symbol of strength is a woman. The Goddess Shakti."

Bhim Rao Ambedkar

Dr. Bhimrao Ramji Ambedkar, was one of the most learned Indian politicians of his time. He yearned to learn, but surprisingly he wanted to stop studying at the age of 10. Nevertheless, as we all know, the story turned out differently.

He was born in a backward caste and faced discrimination. Even in school, he was humiliated, and treated as an ‘untouchable’ because of his caste. He was not even allowed to touch the tap for drinking water. When he was really thirsty, the peon would pour water into his upturned mouth, taking care not to touch his body. When the peon was unavailable, he had to stay thirsty all day long.

Once, he thought of running away from school, and taking a job in a factory in Bombay, but decided against it. Trying to run away was useless. He decided to face his problems. He studied hard to gain knowledge and respect. This self-motivation proved to be a life-changing event not only for himself, but also for the nation.

After he graduated from Bombay University, He received a scholarship to go to Columbia University and The London School of Economics. When he came back, he took it upon himself to end discrimination in India. He became an economist and a social reformer, who inspired the Dalit Buddhist movement. Later, he was appointed by the Assembly to write Independent India’s first Constitution.

“They tried to bury me. They did not know that I was a seed.”

CHAPTER 4

Does Health Care Save Battered Women's Lives: Evidence from the U.S.

Abstract: This study examines the effect of medical care on domestic violence. I find that improving access to medical care significantly decreases domestic violence against women, and have no effect on domestic violence against men. I corroborate this by instrumenting private health insurance coverage with premium taxes. I close by examining the effect of expanding Medicaid coverage. Evidence indicates that the adoption of the ACA Medicaid expansion leads to a 32% decline in domestic homicides against women. Exploration of mechanisms suggests that the Medicaid expansion can lower domestic violence by reducing women's economic dependency on their spouses.

> *And yet the same question, "Why doesn't she leave him?" or its obverse, "Why does she stay?" continues to gnaw at the moorings of the domestic violence revolution. The durability of abusive relationships remains their central paradox......*
>
> Evan Stark,
>
> *Coercive Control: The Entrapment of Women in Personal Life* (2009).

4.1 Introduction

In the United States, around 36 percent of women under age 65 experience financial burdens of health care (Centers for Disease Control, 2012).[1] As adverse physical, mental, and sexual and reproductive health consequences of domestic battery can result in a great need for health care, battered women face much heavier financial burdens than non-battered women (García-Moreno et al., 2015).[2] The economic barrier to health care for them continues to be one of the major policy concerns (Daniel et al., 2018).

What are the effects of increasing health insurance coverage to battered women on victimization? On the one hand, around a third of women are covered as dependents under their husbands' insurance policies (Current Population Census, 2019). Offering a woman's independent option can decrease their economic dependency on spouses, and therefore help them to leave abusive relationships. On the other hand, empowering women economically may lead to backlash from their spouses (Bobonis et al., 2013; Field et al., 2016; Erten and Keskin, 2018), potentially even escalating the level of stress or conflict in intimate relationships.[3]

[1]Women incur greater health care costs than men, and they are more likely to skip a recommended health test or treatment due to cost (Ranji et al., 2019).

[2]The average health costs by injury type range from \$1,700 to \$95,300 in the United States (Peterson et al., 2021).

[3]Women's access to health care may also trigger the police interventions. Such interventions might deprive abusers of their "rightful authority or control" in partners, escalating the level of stress or conflict in intimate relationships. For example, a physician in California documented a retaliation case: "A woman came into an Emergency Department in Los Angeles with facial injuries. A mandatory report was made to the police, resulting in her husband being arrested in the hospital

Although domestic violence has recently gathered attention from economists, there is little definitive evidence of whether places with and without affordable health care vary in their experience of domestic violence. This stands in contrast to other arenas such as health outcomes as a result of biological differences, where large literature documents differences in women's pregnancy and menopause, and gynecological conditions arising from seeking health care services (Weissman and Olfson, 1995; Bennett, 2002; Combs Thorsen et al., 2012; Espey et al., 2019). The existing studies that do relate health care expansion to crime focus exclusively on overall crime (Wen et al., 2017; He and Barkowski, 2020; Borgschulte and Vogler, 2020) and are also difficult to interpret, since the underlying mechanisms for violent and property crime could be different.

This paper provides the first empirical evidence on the effect of women's access to health care on domestic violence. I approach estimation in three phases. In the first, I study variation in medical care prices and women's private insurance coverage, while in the second I concentrate on variation in access to health care exogenously due to differences in insurance premium tax across states. My initial models include fixed effects, time-varying local controls, and local linear time trends to control for a large range of potentially important confounding factors. My analysis of premium tax yields important insights into the implications of levying taxes on insurance premiums for access to health care that are less understood.[4] In fact, only a few states have changed the premium tax rate over the last several decades (Grace et al., 2014; National Association of Insurance Commissioners, 2017). Thus, the premium tax does not respond to a sharp fall in domestic violence. Using both OLS and IV approaches enables me to offer stronger causal evidence about the impact of affordable health care.

waiting room. After the patient was finished being treated for her injuries, she went home thinking she was safe. However, her husband had already been released and was home waiting for her. She came back to the Emergency Department later that day "worse off" than she had been on the first visit as a result of her husband's retaliation"(Hyman, 1997).

[4]The reason that I focus on private insurance coverage in the first two phases is that people insured through public insurance program such as Medicaid are exempt from premiums or eligible for premium tax credits in most states (Centers for Medicare and Medicaid Services, 2014; Blavin et al., 2018).

In the third, I explore the impact of Medicaid expansion after the Affordable Care Act on domestic violence.[5] Medicaid is the largest public health insurance program, which covers 1 in 5 Americans. Prior to the passage of the Affordable Care Act (ACA), Medicaid was mainly used to assist only very poor people. The ACA expands Medicaid to all adults with income up to 138% federal poverty level. The Medicaid expansion was made a state option, while 24 states and Washington DC adopted expansion on January 1st, 2014, the first day of ACA eligibility expansions. Leveraging a dynamic difference-in-difference design, I exploit variation in the state and timing of the Medicaid expansion to compare changes in domestic violence in expanding states to domestic violence in non-expansion states. In doing so, this paper can paint a complete picture of whether or not health care affects domestic violence.

I find that improving women's access to health care decreases domestic violence. This is true under both the direct approach based on my primary measures and the approach based on premium taxes. Lower medical care prices or higher private insurance coverage in a place can save battered women's lives in that place. Also, I find that states that expand Medicaid experience a reduction in domestic violence relative to nonexpansion states. These findings are economically important, as they suggest that making health care more affordable and accessible is an effective policy lever to support battered women, which in turn, can cost-effectively save lives.

The second focus of this study is to understand the mechanism behind the effect of health insurance on domestic violence. One potential mechanism is that the Medicaid expansion can provide abused women with an independent health insurance and reduce their economic dependence on spouses. I find considerable evidence that the ACA Medicaid expansion decreases women's dependent coverage under their spouses' health insurance plans. Also, evidence suggest that domestic violence is reduced, entirely in states experience an above median decrease of women's dependent coverage after the insurance expansion. Therefore, I inter-

[5] In 2019, percentages of women under age 65 with private coverage and public coverage are 64 and 27, respectively, in the United States (Centers for Disease Control, 2019).

pret these results as reflecting the direct impact of empowering women via offering independent and secure health insurance plans in decreasing domestic violence.

Is the decrease in domestic homicides driven by the deterrence effect? In most states, health care providers have a legal obligation to report injuries or domestic abuse to the police. One of the main goal of mandatory reporting is holding perpetrators accountable (Hyman et al., 1995; Lizdas et al., 2019). In this regard, seeking out health care may trigger a response by the criminal justice system, and deter potential domestic violence. I examine whether there exist heterogeneous effects based on the size of policing. The results rule out the possibility that the decline in domestic violence is driven by deterrence.

The natural question to ask is whether insurance expansion has generalized effects on violence against women, regardless of crime types. If so, different potential channels might be in operation. To tackle this concern, I perform falsification checks by investigating its effects on non-domestic violence against women. There are no changes after improving access to the Medicaid program. This finding is useful to reveal that lack of secure health insurance is an issue, particularly for battered women.

In broad terms, I see my work fitting into three main strands in the literature. First, it closely relates to the economics of crime literature on domestic violence by examining the role of health services for battered women. Existing research explores the effects of gender pay gaps (Aizer, 2010), labor market conditions (Anderberg et al., 2016; Bhalotra et al., 2020), family types (Tur-Prats, 2019), emotional cues (Card and Dahl, 2011), intergenerational transmissions (Pollak, 2004), and ancestral features (Alesina et al., 2016).

Second, this paper contributes to the literature on the impact of increasing insurance coverage on crime outcomes. These studies have shown the effect of the Medicaid expansion induced by the Health Insurance Flexibility and Accountability initiative (Wen et al., 2017) and the Affordable Care Act (He and Barkowski, 2020; Borgschulte and Vogler, 2020). However, unlike that literature, which is primarily concerned with public insurance coverage, this paper considers both public and private insurance coverage. That literature also concentrates on overall crime as the outcome of interest, while my outcome is domestic violence.

Third, I view this paper as additionally linked to growing studies of female-friendly services. The bulk of this literature looks at supportive services for female victims of domestic abuse including domestic abuse shelters (Schechter, 2021; Sims, 2021), women's police stations (Amaral et al., 2021), women's justice centers (Kavanaugh et al., 2018), and nutritional assistance (Carr and Packham, 2021)[6]. This study contributes to the literature by shining light on the importance of health care services in reducing domestic violence.

The remainder of the article is structured as follows. Section 2 describes the data and offers some descriptive analysis. Section 3 reports empirical strategies that are used. Section 4 conducts the analysis of medical price and private insurance coverage. Section 5 does the same for the Medicaid expansion. Section 6 concludes.

4.2 Data and Initial Descriptive Analysis

My analysis matches data on crime to prices and insurance coverage, with the aim of estimating elasticities for domestic violence with respect to health care. This is done in three main ways. First, sources of crime data are matched to medical care prices and private insurance coverage. Second, in order to identify the causal effects, state insurance premium taxes are studied. Third, in the case of public insurance coverage, a dichotomous indicator is created, which captures the implementation of Medicaid expansions.

4.2.1 Medical Care Prices

The data source for consumer price index (CPI) is the FRED, Federal Reserve Bank of St. Louis. It measures inflation by tracking retail prices of a good or service of a constant quality and quantity over time. The major component of medical care

[6]Parallel studies also focus on the effect of female police officers (Miller and Segal, 2019) and female jurors (Anwar et al., 2019)

index is professional services. The data are available for 27 metropolitan statistics areas (MSAs) in the United States. I collect annual prices data between 1976 and 2019, which can be matched to homicides data.

4.2.2 Private Insurance Coverage

Data on private insurance coverage come from the Current Population Survey. This survey is one of the largest and most well-recognized surveys in the United States. Participants are asked questions for their insurance coverage, including whether they were covered by a private insurance plan during the preceding calendar year. Having collected the data for surveys between 2017 and 2019, I calculate the average private insurance rate for each state that is used to match to state premium tax rate. In order to perform the falsification analysis of the effect of premium tax on private insurance coverage, I collect additional data on public insurance and compute coverage rates for Medicare and Medicaid, respectively, in the same vein.

4.2.3 Premium Taxes

Premiums collected by insurance companies have been subject to state taxation in the United States. The state premium taxes are often a percentage of the premiums paid by the insured. I obtain data on health premium tax rates for states and Washington DC from National Association of Insurance Commissioners (NAIC), which are available only for 2017 and later. Over the period, none of states change their policy on taxation. Oregon and Wisconsin are excluded in the main analysis, as they impose taxes in different ways (such as imposing fixed taxes).

4.2.4 Medicaid Expansion

Medicaid is a publicly funded health insurance program for low-income families and individuals in the US, which is founded in 1965. Prior to the Affordable Care Act (ACA), one consistent policy has been categorical restrictions that limit coverage to the disabled, children, and members of families with dependent children.

The ACA expands Medicaid benefits to all adults with incomes at or below 138 percent of the federal poverty level (FPL). Many states have started to implement the ACA Medicaid reform since 2014.

I collect data from the Kaiser Family Foundation (2019c) to identify states that expand Medicaid and the dates that the expansions become effective. Table 1 presents the names of states that expand Medicaid through the ACA, and the corresponding dates. Then, they are matched with state-level crime data for the years 2009 through 2019. My main analysis is at the state and month levels. I classify the expansion treatment for New Hampshire to begin in August of 2015 because the reform becomes effective in the first half of August. While for analysis at the state and year levels, I code treatment to start in 2015, 2016 and 2017 for New Hampshire, Alaska, and Louisiana, respectively, as these expansions do not become effective until late 2014, 2015, and 2016.

4.2.5 Domestic Homicides

The homicides data are collected from the Supplemental Homicide Reports (SHR) within the Uniform Crime Reporting (UCR) program, starting in 1976. The SHR contains detailed information on relationships between victims and perpetrators and victim sex. Domestic homicides include homicides committed by a current or former boyfriends, girlfriend, or spouse of a victim.

In the primary analysis I categorize domestic homicides into two types: (i) the victim is female, (ii) the victim is male. The main outcome variables are domestic homicide rates by victim sex at the MSA or state level, calculated by dividing domestic homicide counts respectively by the population (in 100,000s) in that MSA or state.[7]

[7] Data on homicides for Alabama and Florida are not available for the Medicaid expansion analysis at the state-month level.

4.2.6 Other Characteristics

To take account of the effect of policing, I use the information on police employment obtained from the Law Enforcement Officers Killed or Assaulted (LEOKA) collection. I supplement my analysis with additional data on social-economic characteristics including income, unemployment, and population from the Current Population Survey and U.S. Census.

4.2.7 Initial Descriptive Analysis

Figure 1 shows trends in the total number of domestic homicides by victim sex in the United States. What stands out is the decline of domestic homicides, regardless of victim sex. Domestic homicides against women (the left axis) fall from just under 1,200 in 1976 to just over 900 in 2019, while domestic homicides against men (the right axis) drop from just over 900 in 1976 to just under 300 in 2019.

Figure 2 shows average annual changes in the medical price index and the domestic homicide rates for female victims and male victims, respectively. The averaged changes in domestic homicides against women and medical prices are evidently associated as the positively sloped (and statistically significant) regression line fit through the points in Panel A. On the contrary, such a relationship is not clear for domestic homicides against men. Study of the two charts indicates that economic burden of health care might disproportionately affect battered women. In addition, Table 2 displays summary statistics of key variables used in the main analysis for analysis of Medical prices and private insurance.

In terms of the analysis of the Medicaid expansion, Table 3 presents population weighted means and standard deviations of baseline variables,measured for 2009–2013, in expansion and non-expansion states, respectively. Expansion states generally have experienced slightly less domestic violence against women. In terms of other baseline characteristics such as labor market conditions for women and men, they are very similar.

4.3 Empirical Strategy

4.3.1 Baseline Model: Medical Care Price

The primary outcome variable is the domestic homicide rate because it is a measure of the most severe form of violence and a major policy concern. Moreover, endogenous reporting is unlikely to be a concern for homicides. My empirical specification for the effect of medical care prices is specified as follows:

$$Y_{mt} = \beta_1 MedicalCarePrice_{mt} + \beta_2 X_{mt} + \sigma_m + \phi_t + m_t + \varepsilon_{mt} \tag{1}$$

Where Y_{mt} is the rate of domestic homicide per 100,000 population in MSA m in year t. My main focus is on domestic abuse against women but, as a falsification exercise, I examine domestic abuse against men. The main explanatory variable, $MedicalCarePrice_{mt}$ is the price index for medical care in that locality and year. The term X_{mt} is a set of controls for all goods price, food price, alcohol price, police employment by gender, unemployment by gender, average income by gender, and population. σ_m and ϕ_t are MSA and year fixed effects, respectively. To account for MSA-specific linear time trends, I also include the trend control, m_t. The term ε_{mt} represents the error term.

4.3.2 Baseline Model: Private Insurance Coverage

I also examine the impact of private insurance coverage on domestic violence, as private insurers pay the largest share of patients' medical costs (Centers for Disease Control, 2018a). I estimate the following equation:

$$Y_s = \beta_1 PrivateInsurance_s + \beta_2 X_s + \varepsilon_s \tag{2}$$

Where Y_s is the mean rate of domestic homicide per 100,000 population in state s between 2017 and 2019. $PrivateInsurance_s$ is the mean rate of women's private insurance coverage in that state. As before, X_s includes controls such as police employment by gender, unemployment by gender, average income by gender, and population.

4.3.3 Instrumental Variable Approach

Although the above frameworks can deal with a large range of possible confounding factors, they may not have entirely solved the potential issue of omitted variables. This motivates my use of instrumental variable approach. For a valid instrument variable, it influences domestic homicides against women exclusively through women's private insurance coverage. I use state health insurance premium tax rates to instrument women's private coverage rates. As discussed, there is a dramatic change in domestic homicides against women over the past several decades, while premium tax rates are not adjusted in response to such a change. In this regard, this approach can be considered as a quasi-experiment where variation in premium tax rate across states is exogenous.

The main reason for women having difficulties purchasing insurance is high premiums (Centers for Disease Control, 2015).[8] Therefore, the relevance condition for a valid IV is satisfied, as clearly the instrument is correlated with the potentially endogenous regressor. Also, the results of the "horse race" analysis support the exclusion restriction condition.[9]

According to Imbens and Angrist (1994), the standard IV assumption would be violated if the instrument is related to domestic violence through some variable other than private insurance coverage. For instance, the assumption would not hold, if the premium tax affects materially economic welfare for women, and thus domestic violence.

The IV assumption seems plausible. First, the premium tax is a small amount of revenue for government. In 2021, the revenue from insurance premium tax is less than $25 million in the US, which only accounts for around 2% of total tax rev-

[8]In 2020, the average monthly premium of health insurance in the United States is $505 (Price, 2020).

[9]See Section 4.3 for details.

enue (Census Bureau, 2021). Second, improving women's welfare requires large public expenditures. Based on the estimation of National Women's Law Center (2022), millions of women are still struggling to make ends meet in the US. The US government has announced an historic commitment of over $12 billion dollars to improve women's welfare that includes advancing women's employment and closing gender pay gap (White House, 2021). A small amount of revenue from insurance premium tax is unlikely to play an important role in improving women's welfare. Third, I conduct further exercises where the women's economic outcomes are used as the dependent variables. Table A4.2 shows the estimates of the effect of premium tax rate on female employment and income. The results suggest no significant association between premium tax and women's economic welfare, offering support for its plausibility. Lastly, my horse race analysis also provides suggestive evidence in favour of the assumption in Table 8.

The IV approach is to identify the local average treatment effect (LATE), that is, the average treatment effect on the complier subpopulation (Imbens and Angrist, 1994). In my IV setting, the LATE applies to those women whose private insurance uptake is affected by the insurance premium tax. This group is not small. In the US, there are 35 million uninsured nonelderly adults and 70% of them are workers in 2013. More than 60% of these workers are employed by small companies with fewer than 100 employees (Kaiser Family Foundation, 2014). Less than half of workers are offered health insurance through their employers (Robertson et al., 2012). The cost of health insurance, particularly premium, is an important factor for insurance coverage among small businesses (Buettgens et al., 2010; Gruber and Lettau, 2004). Therefore, small firms could be sensitive to the insurance premium tax, which could affect private insurance coverage (employer-provided coverage accounts for 90% of total private coverage). For instance, Jensen and Gabel (1992) find that the probability of coverage is extremely elastic to the insurance premium tax rate: one unit increase in the average tax rate would cause the probability of private insurance coverage to fall from 0.488 to 0.323.

In the separate regression, I test for effect of premium tax rates on domestic homicides against men, finding no effect. This result shuts down any unobservable channels through which premium taxes influence overall domestic violence. As already mentioned, premiums are not required for public insurance programs such as Medicare and Medicaid in most states. That is why private insurance coverage is used in the main analysis. In the Appendix Table A4.1, I present evidence suggesting that premium taxes do not influence Medicare and Medicaid coverage for women, hence further justifying the use of premium taxes as an instrument for women's private insurance coverage. I also test whether there is a weak instrument problem, further ensuring the validity of my proposed instrument.

The specifications used to conduct the IV analysis are as follows:

$$Y_s = \theta_1 PremiumTax_s + \delta_1 X_s + \varepsilon_s \tag{3}$$

$$Private\widehat{Insur}ance_s = \theta_2 PremiumTax_s + \delta_2 X_s + \varepsilon_s \tag{4}$$

$$Y_s = \theta_3 Private\widehat{Insur}ance_s + \delta_3 X_s + \varepsilon_s \tag{5}$$

Equation (3) presents the reduced form specification where $PremiumTax_s$ is the health premium tax rate in state s. Equation (4) is the first-stage equation that associates premium tax with women's private insurance coverage. Equation (5) corresponds to the second-stage equation. The IV estimate is then the ratio of the reduced form to the first stage coefficient, so that θ_3=θ_1/θ_2.

4.4 Results

4.4.1 Baseline Results: Medical Care Price

This section describes estimates from Equation (1), which examines the effect of medical care prices. Table 4 shows the impact of medical care prices on domestic homicides against women. Columns 1 through 4 report results without controls, with all goods price control, with full controls, and with full controls plus MSA-specific linear time trend controls, respectively. Across specifications, the estimates are statistically significant, strongly implying that the estimated effect is thus insensitive to the inclusion of control variables.

The magnitudes of the estimates are nearly identical through the varying specifications, ranging from 0.06 to 0.08. The preferred specification is displayed in Column 4 with MSA-specific linear time trends. Identification of the impact of medical care prices in such a specification comes from within-MSA variation after netting out variation in domestic homicides against women caused by factors that vary linearly over time and that are specific to individual MSAs. The point estimate of 0.07 indicates that increasing medical care price index by 0.26 (equivalent to around 10% of mean value) increases the female domestic homicide rate by around 0.02 homicides per 100,000 population (equivalent to around 5% of mean value).

Table 5 presents corresponding estimates for domestic homicides against men. Unlike estimates in Table 4, these estimates are highly unstable. Specification 1 estimates the effect of medical care prices on domestic homicides against men without controls, suggesting a positive and significant effect. However, once eliminating variation in such homicides driven by the overall price index in specification 2, the point estimate not only loses statistical significance but also drops considerably in magnitude. It suggests that a large part of the medical care price captures is an overall goods price in specification 1. Specification 3 and 4 show that the estimate continues to remain statistically insignificant after accounting for other covariates and MSA-specific time trends. This may be interpreted as implying that the scarcity of medical care is less likely to be a concern for battered man.

4.4.2 Baseline Results: Private Insurance Coverage

The analysis so far has used the medical care price to measure access to health care. Because the primary source of payment for health care resulting from domestic abuse is private insurance, I also use a different proxy of access to health care, private insurance coverage, to investigate whether the results from the two measures are in parallel.[10]

Figure 3 depicts a strong, negative and significant correlation between average private insurance coverage rates of women and average female domestic homicide rates between 2017 and 2019.[11] The chart very clearly shows that states with higher private insurance coverage rate — like Minnesota, New Hampshire, and Massachusetts see relative higher rate of domestic abuse against women. On the other hand those with far less coverage — like New Mexico, Louisiana, and Mississippi experience more violence against women.

Table 6 presents estimates from Equation (2), which examines the effect of women's private insurance coverage on domestic violence. Column 1 reveals the unconditional cross-state relationship. Specifically, the estimated coefficient, statistically significant at the 1 percent level, implies that a 10% increase in women's private insurance coverage rate is associated with a decline in domestic homicides against women by around 0.10 homicides per 100,000 population. It is economically meaningful, when compared to mean value of 0.36 homicides per 100,000 population. Column 2 shows that the coefficient retains its sign and continues to remain highly statistically significant after accounting for the influence of law enforcement. Reassuringly, the estimate in Column 3 is robust to the introduction of other important confounding factors.

Overall, there are two points of note. First, the baseline analyses are based on different measures of women's access to health care, yet both show a significant elasticity for domestic violence with respect to health care. As a result, these findings lend credence to the assertion that health care can save battered women's lives. Second, the absence of a negative (positive) impact of medical care price (insurance coverage) on domestic violence allows me to interpret these results as evidence that there is no significant backlash against women in response to their access to health care.

[10] Appendix Table A4.3 displays the distribution of primary source of payment for health care resulting from domestic violence against women.

[11] There are 49 states and Washington DC in the analysis. Data on homicides for Florida are not available during the time period.

4.4.3 Instrumental Variable Results

The results presented in the baseline analysis give me confidence that a robust and strong relationship exists between health care and domestic violence. However, a more rigorous evaluation is required to tackle the potential issue of omitted variables. A quasi-experimental design can be set up where variation in private insurance coverage across states is explained by the health insurance premium tax policy that did not change in response to a dramatic change in domestic homicides.

Figure 4 provides support for the relevance condition with respect to a valid IV by showing that premium tax is statistically significantly correlated with women's private insurance coverage. The descriptive statistics for health insurance premium tax are displayed in Table 7. An evident feature of the Table is that a large number of states impose a tax rate in the range of 1% - 2%, whereas only few states are under or over the range. Columns 1 and 2 present the average rate of women's private insurance coverage and domestic homicides against women by ranges of premium tax rates. Importantly, inspection of Column 1 further suggests that the higher a premium tax rate is, the lower a private insurance coverage rate. Turning to domestic homicides against women in Column 2, there are positive and negative associations with premium tax and insurance coverage, respectively.

For the exclusion restriction condition, I conduct a formal examination of whether premium tax influences domestic violence solely via the effect of private insurance coverage. This is a particularly important examination because, if premium tax influences domestic violence either directly or via some other unobserved channels, then the analysis using the IV approach would be attributing this latent effect to the private insurance channel.

Building on the work of Ashraf and Galor (2013), I implement the aforementioned examination by the "horse race" analysis. First, I estimate a specification that includes premium tax rather than private insurance coverage to explain the cross-state variation in domestic homicides against women. Second, I examine an additional specification that includes both premium tax and private insurance coverage as covariates. Lastly, the observed results from the first specification are compared to those from the latter. Unless premium tax and private insurance coverage are ultimate and proximate determining factors within the same mechanism, then private insurance, once included in the specification, should not capture most of the explanatory power otherwise attributed to premium tax.

The corresponding results are reported in Table 8. Column 1 shows the estimate of the unconditional effect of women's private insurance coverage on domestic homicides against women, which is statistically significant. Even though Column 2 shows a highly statistically significant unconditionally impact of premium tax on domestic homicides against women, however, such an impact not only lose statistic significant but also drops substantially in magnitude after controlling for women's private insurance coverage in Column 3. Strikingly, conditional on premium tax, the estimate associated with private insurance remains highly statistically significant, reflecting that private insurance survives a horse race. Taken together, these results provide strong evidence in support of the view that premium tax affects domestic homicides against women via women's private insurance coverage alone.

Table 9 shows estimates of four specifications given by Equations (2)-(5), namely the OLS estimate, the reduced form estimate where the instrument is the premium tax, the first stage estimate, and the IV estimate. Considering the OLS estimate first, it is clear that there is significant and negative elasticity of domestic violence with respect to private insurance coverage in Column 1. However, the OLS estimate may be downward biased — for example, if the abusers prevent victims from accessing health care even when the associated costs are covered.

Further to this, there is a strong reduced form relationship between domestic violence and premium tax (see Column 2). The magnitude of the estimate indicates that a 1 percentage point increase in health premium tax rate (equivalent to 50% of mean value) leads to an increase of 0.07 deaths per 100,000 population (equivalent to 19% of mean value). The first stage in Column 3 is also strong with an F-statistic of 18.14, and the point estimate suggests that a 1 percentage point increase in health premium tax rate (equivalent to 50% of mean value) decreases women's private insurance coverage rate by 3 percentage points (a 5% decrease relative to the mean value).

Turning to the IV estimate, Column 4 of Table 9 confirms a negative and statistically significant effect of women's insurance coverage on domestic violence. More specifically, the coefficient of -2.07 shows that a 5 percentage point increase in women's private insurance coverage rate (corresponding to an increase of 8% of mean rate) decreases domestic homicides against women by around 0.10 homicides per 100,000 population (27% of mean rate). Put differently, it implies that 6 battered women's lives can be saved in a state, if there are around 170,000 new women enrolling a private insurance in that state. The larger coefficient on the IV estimate relative to the OLS estimate is consistent with downward endogeneity bias on the OLS estimate.

As a falsification check, I examine the effect of premium tax on domestic violence against men and find no effect in Table 10. This result is consistent with the claim that access to health care is a concern mainly for battered women. Moreover, it rules out any potential explanations of generalized effects of premium tax on domestic violence against women and men at large.

Lastly, I explore whether there are generalized effects that reduce violence against women due to improving access to health care. Access to health care and domestic violence are very closely linked through repeat victimization but the same may not be true for violence against women in general. The same exercises as Table 9 are repeated using homicides against women as the dependent variable. The results in Table 11 suggest that health care has no effect on violence against women. Hence, my results are unlikely to be driven by generalized mechanisms.

Homicides have high direct costs to society. Cohen and Piquero (2009) estimate that the cost per homicide offence is around $7 million. Back of the envelope calculations suggest that a 1 percentage point decline in premium tax rate can save 4 women's lives per year in a state, and thus yield an average saving of $28 million. Given that the average premium tax revenue for one state is $0.5 million in 2019, improving insurance coverage for women by reducing taxes for premium has the potential to result in large savings.

4.5 Medicaid Expansion

Thus far the analysis has focused on the crime-reduction effect of private health insurance coverage, but expanding public health insurance program may also impact domestic violence. In the United States, Medicaid is the largest public health insurance program, which covers 1 in 5 Americans. Prior to the passage of the Affordable Care Act (ACA), Medicaid eligibility was tied to assistance for only very poor people in specific categories (e.g. children and some of their parents, disabled people, and pregnant women). The ACA expands Medicaid to nonelderly adults with income up to 138% federal poverty level, which increased average annual enrollment by 6.3 million (Centers for Medicare and Medicaid Services, 2015). The Medicaid expansion was made a state option, so it allows to examine the effects of the ACA Medicaid expansion on domestic violence by exploring variation in the state and timing of the expansion. Figure 5 indicates the states that adopt the Medicaid expansions by the end of 2019.

4.5.1 Empirical Strategy

The primary outcome variable is the domestic homicide rate because it is a measure of the most severe form of violence and a major policy concern. Moreover, endogenous reporting is unlikely to be a concern for homicides. To estimate the effect of the ACA Medicaid expansion on such a measure of domestic violence, I exploit the variation in expansion placement and timing applying a flexible difference-in-difference (DD) specification. The model enables me to include state fixed effects to take level difference between states into account and time fixed effect to control for unobserved shocks.

I analyze whether the ACA Medicaid expansion affects domestic homicides using the following base equation:

$$Y_{st} = \beta MedicaidExpansion_{st} + \sigma_s + \phi_t + \varepsilon_{st} \tag{6}$$

The dependent variable Y_{st} is the number of domestic homicides in state s in month-year period t ($t = my$, where m is month and y is year). The variable $MedicaidExpansion_{st}$ is an indicator that takes the value 1 for the post-expansion period, and 0 otherwise. The terms σ_s and ϕ_t correspond to indicators for state and time, respectively. Drawing on Bertrand et al. (2004), all SEs are clustered at the state level to allow for correlation of errors over time within states.

The identification strategy compares the post-expansion changes in domestic homicides in states that implement the expansion with changes in those that do not implement the expansion. Two-way fixed effects DD estimates will yield a biased estimate of the causal effect of the ACA Medicaid if expansion states would have experienced a decrease in domestic homicides even absent the expansions. I cannot rule out such a scenario, but I can test for differential trends in domestic homicides between treatment and control states in the years leading up to expansions.

Furthermore, a recent body of research has investigated the biases of the two-way fixed effects DD estimator associated with heterogeneous treatment effects (Sun and Abraham, 2021; Callaway and Sant'Anna, 2021; Athey and Imbens, 2021; Goodman-Bacon, 2021). In particular, such an estimator equals a weighted average of all possible simple 2x2 DDs that compare one group that changes treatment status to another group that does not. In other words, earlier-expanding states are compared with later-expanding states, where earlier-expanding states serve as controls, which might lead to a bias in presence of heterogeneous treatment effects across states adopting Medicaid expansion at different time points.

To tackle this concern, I follow Sun and Abraham (2021), who develop DID estimators of the cohort-and-period specific effects that only rely on the parallel trends assumption, and that are robust to heterogeneous treatment effects. Their estimators categories units into different cohorts based on their initial treatment timing, to avoid the estimates of lags and leads being contaminated by effects from other periods. Specifically, they define their estimates as "interaction-weighted" estimates, which are generated mainly in two steps. First, the cohort average treatment effects on the treated (CATT) are computed by estimating the cohort-specific average difference in outcomes relative to never being treated. Second, their estimator estimates a weighted average of CATT with weights equal to the share of each cohort in the relevant period(s).[12] Based on their methodology, I estimate the following equation:

$$Y_{st} = \sum_{e=1}^{E} \sum_{\substack{k=-5 \\ k \neq -1}}^{4} \beta_{e,k} (\mathbb{1}\{E_s = e\} MedicaidExpansion_{st}^{k}) + \sigma_s + \phi_t + \varepsilon_{st} \tag{7}$$

Where time windows span periods of one year each. Particularly, $\pm k$ ranges from 4 and -5 to respectively 1 and -2 years, as -1 is omitted. Each lag(lead) takes the value of the main regressor -k(k) years away from the ACA Medicaid expansion. $MedicaidExpansion_{st}^{k}$ is a set of relative event-time dummies, that take value of 1 if time t is k years after (or before, if k is negative) the expansion. e represents cohorts, different time periods in which states implement the expansion. Each estimated parameter is a weighted average of the impact of the expansion k years away (i.e., pre or post). The coefficients of interest $\beta_{e,k}$ can be interpreted as an average effect of the treatment on the treated periods after initial treatment.

4.5.2 Results

In this section, I show that expanding health insurance coverage can reduce domestic violence. I start by presenting evidence on the effects of the ACA Medicaid expansion on domestic homicides against women. Commencing with the two-way fixed effects DD and dynamic DD results on the number domestic homicides, I later assess the impact on extensive margin of domestic homicides. A series of sensitivity

[12]Their estimators use the never-expansion units are used as controls if there are never-treated units. Their procedure ensures nonnegative weights and better sheds light on dynamic treatment effects.

checks are performed to ensure the robustness of the findings.

4.5.2.1 Main Results

I first examine the effects of the ACA Medicaid expansion on domestic homicides against women by estimating Equation 7. Table 12 presents the results. Column 1 shows the estimate including state, year, and month fixed effects. Column 2 replace year and month fixed effects with year-month fixed effects. Column 3 adds other controls including the log of police rate, unemployment by gender,and income by gender at the state-year level. Column 4 presents the result with the inclusion of state-specific time trends.[13] Throughout the specifications the coefficients are statistically significant and negative, suggesting that the ACA Medicaid expansion is associated with less domestic violence against women. In particular, the estimate in Column 4 implies that expanding Medicaid decreases domestic homicides against women by approximately 0.55 homicides (equivalent to 16% of mean value).

As a falsification check, I examine the effect on domestic violence against men. Appendix Table A4.4 repeats the same exercise using domestic homicides against men as the outcome variable. None of the results in Columns 1-4 is statistically significant, indicating that unmet need for health insurance is a concern mainly for battered women. Moreover, it rules out any potential explanations of effects of the expansion on domestic violence against women and men at large.

Adults are the target of the ACA Medicaid expansion. Using information on victim age, I next explore the effects on the occurrence of domestic homicides against adults and non-adults women, respectively. Table 13 follows suit, displaying the estimates of the effect of the expansion on domestic homicides against adult women. The estimates for adult women, both in terms of magnitude and significance, are strikingly similar to those for all women. On the contrary, I find little evidence that domestic homicides against non-adult women response to the ACA Medicaid expansion (see Table 14). In this light, the findings suggest that providing health

[13]Since it is plausible that state population may change throughout the time, all regressions control for the log of annual population.

insurance to adult victims has no spillover effect on non-adult victims. Also, the absence of a positive impact allows me to interpret these results as evidence that there is no significant backlash against abused women in response to improving access to health care.

I further presents evidence of the validity of my empirical design and dynamic effects from the event study analysis. Figure 6 shows the results of running Equation 7, where the dependent variable is the number of domestic homicides against women, and supports the identification assumption. Specifically, it shows that the coefficients for years prior to the ACA Medicaid expansion never reach statistical significance, even at the 10% level, indicating little evidence of differential group trends. Importantly, the Medicaid expansion marks a clear break from the pattern of the pre-period. Following expansions, domestic homicides decrease significantly. Domestic homicides decline by around 0.61 homicides in the year of the expansion and by 0.59 in the following year. Effects then dissipate, reaching insignificance two years after expanding.

A number of factors might explain the pattern of effects. First, health insurance plans in the ACA marketplace are unaffordable for many middle-class people, who don't qualify for Medicaid or federal subsidies to help buy an individual policy. According to the Kaiser Family Foundation (2019a,b), the average premium for a midlevel plan for a 40-year-old has climbed to \$478 a month in 2019 from \$273 in 2014.[14] Therefore, as relative time increases, the decline in coverage of middle-income adults may offset the rise in coverage of low-income adults.

Alternatively, even though there are additional federal matching funds for expanding Medicaid eligibility, state budgets may be affected. That is, expanding states may have to cut spending in other parts of their Medicaid programs, or cut spending outside of Medicaid (Ward, 2020). The fiscal stress faced by state governments may become more salient when considering periods farther after expanding Medicaid under the ACA. Finally, the ACA was attacked under the Trump admin-

[14]New regulations restricting insurers' ability to charge higher premiums to older and sicker adults may lead to unnecessarily high premiums for younger and healthier individuals (Eibner and Nowak, 2016).

istration, which might also help explain the fading effect of the ACA Medicaid expansion.

I estimate separate event studies for domestic homicides against adult women and non-adult women. As shown in Figure 7, the effects on victimization of adult women for years prior to the Medicaid expansion are relatively small and statistically equal to 0, indicating flat pre-trends. Evidently, the effects occur immediately after extending insurance coverage and last for two years. In contrast, Figure 8 shows that there is no a marked divergence of trends before and after the Medicaid expansion for domestic homicides against non-adult women. These are consistent with the main tables and suggests that adult victims of domestic abuse are exclusively affected.

As an additional robustness check, I test whether the results hold without considering potential biases from heterogeneous treatment effects. Appendix Figures A4.7.1-3 shows the same regressions but employing the two-way fixed effects event study model. The analogous point estimates are documented, suggesting that my findings are robust to varying specifications. More importantly, they serve as implicit evidence of absence of heterogeneous treatment effects across expanding states. Such findings seem reasonable, because most states expanded Medicaid (75%) on the first day of 2014.

4.5.2.2 Extensive Margin: the Probability of Homicide Occurring

An important question to ask is whether expanding Medicaid eligibility brings about a decrease in domestic violence only in states that would have experienced some domestic violence in any case? What about states where the frequency of domestic violence is relatively low? To offer a better understanding of the full picture, I examine the effect of the ACA Medicaid expansion on the probability that at least one domestic homicide against women occurs. The same exercise is repeated for extensive margin and Figure 9 displays the point estimates.

The results show that there seems no change in the pattern of extensive margin of domestic homicides before and after Medicaid expansion, suggesting that the crime-reduction effects of expanding insurance coverage concentrate on violent places. This may be plausibly explained by the fact that the ACA Medicaid expansion does not intend to specifically target domestic abuse victims in the first place. A key implication of this result is that a victim-specific Medicaid expansion may be useful to decrease domestic violence.

4.5.2.3 Placebo Test: Randomization Inference

In light of the concern that the data are highly serially correlated across states, all my regressions have been clustered at the state level. Nevertheless, in this section I conduct a further test to account for this concern. Analogous methods are developed and used in Aglasan et al. (2017), Pinotti (2017), and Ciacci and Sviatschi (2021).

Figure 10 shows the estimates of randomizing the Medicaid expansion treatment across time within states with one thousand permutations. The red vertical line indicates the estimated coefficient in my primary equation with full set of controls, whereas the blue curve indicates the distribution of estimates from the randomization test. If the same estimate can be obtained by chance, I would expect the junction of the red vertical line and the blue curve. The results suggest that the probability of obtaining the identical estimate obtained in my primary equation by change is very unlikely: out of one thousand permutations, none can generate the same result. The result nets out the possibility that serial correlation is biasing my estimate, and again helps bolster validity of my research design.

4.6 Mechanisms

In this section I examine three mechanisms that may help explain the decrease in domestic violence caused by the ACA Medicaid expansion: economic independence channel, deterrence channel and mental health channel. Each of these mechanisms can be explored applying additional data. I also discuss potential generalized channels by directly investigating the effects on non-domestic violence against women.

4.6.0.1 Economic Independence Channel

There is a growing literature devoted to the causes of domestic violence that documents that women's economic empowerment could reduce the prevalence of domestic violence (Aizer, 2010; Bobonis et al., 2013; Anderberg et al., 2016; Tur-Prats, 2019). In particular, Aizer (2010) shows that the way of improving women's outside options involves improving labor market conditions for women that helps terminate abusive relationships and lower levels of domestic violence.

The same may apply to independence of health insurance for women. In the US, around a third of women are covered as dependents under their husbands' insurance policies (CPS, 2019). It is one of the key indicators of women's economic dependency on their partners. As discussed by Short (1998), single women are in a particularly vulnerable position when it comes to health insurance. Unlike their married counterparts, who have spouses to fall back on, single women are less likely to have insurance, and more likely to lose their insurance when they have it. Furthermore, married women who are insured as dependents are less secure than married women with their own health benefits. That is why women with dependent coverage could be more likely to be trapped in risky and violent relationships. In this regard, expanding Medicaid eligibility could reduce women's reliance on dependent coverage, allowing them to make their own choices such as dissolving the relationship.

I approach examination in two ways. In the first, I study variation in dependent coverage for women across states before and after the ACA Medicaid expansion. Specifically, I use data available at the state-year level from the Population Survey. The respondents have been asked if they have coverage, as a dependent, on employment-based health insurance.[15] If the Medicaid expansion leads to greater economic independence for women, I would expect a decline in the proportion of

[15]The majority of Americans (55%) were covered by employer-based health insurance (Berchick et al., 2019).

married women with dependent coverage.[16]

Figure 11 depicts the event-study plots of the effect of the ACA Medicaid expansion on rates of married women with dependent coverage on employment-based insurance at the state-year level, following Sun and Abraham (2021). In the years leading up to the expansion, women's dependent coverage is statistically identical in level and trend between the groups. In the immediate aftermath of the expansion, it decreases sharply among expansion states but remains smooth among those farther away. One may wonder whether the effect is driven by the improvement of labor market conditions for women after the Medicaid expansion. For example, the ACA Medicaid expansions may result in a growth in the nurse labor market. Thus, I explore the effect on women's unemployment rate. Figure 12 shows that coefficients are not statistically different from zero. This finding seems to exclude such a possibility.

In the second, to better understand the salience of this explanation, I explore the heterogeneous effect by the change in women's dependent coverage on domestic homicides. It is important to note that women with dependent insurance coverage may not enroll in Medicaid, though they are eligible for it under the ACA. Particularly, Medicaid may provide less comprehensive services, compared with private insurance programs. To this end, I categorize expansion states into above and below median decrease states based on the percentage decrease of women's dependent coverage in the first year of expansion. Then, estimating Equation 8, I compare them with non-expansion states, respectively.

Figure 13 plots the estimates of the effect of the ACA Medicaid expansion in states with above median decrease in women's dependent coverage on domestic homicides against women. It shows very similar point estimates as in Figure 7. Conversely, Figure 14 shows little evidence of differential trends between expansion states with median decrease in women's dependent coverage and non-expansion states after the ACA Medicaid expansion. Taken together, these findings suggest

[16]Generally, legal spouses are eligible for dependent health insurance (Health Markets, 2019). Hence, my analysis focuses on married women.

that the effect of expanding Medicaid is primarily driven by a decrease in women's reliance on their spouses' health insurance plans. This is again consistent with the economic independence channel through which the Medicaid expansion reduces domestic violence.

4.6.0.2 Deterrence Channel

Heath care providers can help battered women obtain justice even though the first thing that victims of domestic violence need is often medical attention. Crime reporting is a crucial input into the production of law enforcement services. Most U.S. states have enacted mandatory reporting laws that require the reporting of injuries and wounds to law enforcement in health care settings. For example, the law in California requires medical professionals to notify local law enforcement if he or she provides health services for a physical condition to a patient whom he or she knows or reasonably suspects is suffering from any wound or other physical injury that is the result of assaultive or abusive conduct.[17] As a result, improving victims' access to health care may increase the the likelihood of police interventions. This might ultimately increase the expected costs of abusing partners, deterring potential abusers from committing violence. The logic I put forth parallels the deterrence theory that an increase in an offender's chances of being caught decreases crime (Chalfin and McCrary, 2017).

I examine heterogeneity by police size. Specifically, I incorporate state police rates into the regression specifications by interacting the post dummy with the state police rate (and including the police rate as a standalone variable). If deterrence matters, high rates of police would meaningfully reinforce the estimated effect in violence reduction.

Figure 15 shows the corresponding results. In the figure, I assess the effect of the Medicaid expansion on domestic homicides both at a low level of policing (the twenty-fifth percentile: 189 officers per 100,000 persons) and at a high level of polic-

[17] https://www.futureswithoutviolence.org/wp-content/uploads/Compendium-4th-Edition-2019-Final.pdf

ing (the seventy-fifth percentile: 256 officers per 100,000 persons). Moving from the twenty-fifth percentile of police rate to the seventy- fifth percentile does not statistically increase the crime-reduction effect. Hence, this finding is not in favor of the deterrence interpretation.

By and large, these results do not support the deterrence mechanism connected to changes in domestic violence occurring at the same time as expansions of the Medicaid eligibility.

4.6.0.3 Mental Health Channel

In this section, I examine one alternative account positing that the occurrence of domestic violence relates to mental health. There is a significant association between lifetime domestic violence and a wide range of mental health problems in women (Tol et al., 2019). On the other hand, Symptoms of mental illnesses include aggression and agitation, an exaggerated sense of self-confidence, poor decision-making, racing thoughts, and risk-taking behaviors. The Medicaid expansion may bring about an increase in the use of mental health services for potential offenders, which in turn, can reduce domestic violence (analogous to the effects of mental healthcare on crime in Deza et al. (2022)).

To explore it, I test the effect mental illness prevalence by employing data from the U.S. Department of Health and Human Services for 2009-19.[18] Figures 16 and 17. present the point estimates for prevalence of mental illness and serious mental illness, respectively. All posttreatment estimates are insignificant, suggesting that the crime-reduction effect does not seem to arise from such preventive interventions. This result is consistent with the finding in Wen et al. (2017) that treatment for mental illness does not influence homicides.

[18]It publishes data on the prevalence of (serious) mental illness among adults at state and year levels.

4.6.1 Generalized Channels for Non-Domestic Violence

Lastly, I explore whether there are generalized effects that reduce domestic violence due to improving access to health care. As discussed earlier, insurance coverage and domestic violence are very closely linked through economic dependency for spouses in the households but the same may not be true for non-domestic violence against women. The same exercise as Figure 7 is repeated using non-domestic homicides against women as the dependent variable. The results in Figure 18 suggest that the Medicaid expansion has no effect on other violence against women.

The result also provides important insight into the observed effect on domestic violence, because it shows that such an effect is unlikely to be driven by any generalized mechanisms for other crime types. For example, expanding Medicaid insurance program may reduce domestic violence through timely health care for acutely injured victims of domestic violence. If this is the case, I would expect the same effect on non-domestic violence. However, the finding does not support it. Similarly, potential mechanisms such as healthcare facility expansion could be ruled out.

Overall, while no examination individually offers conclusive evidence of the presence or absence of a particular channel, I, on the basis of these results from different tests, interpret the effect of the ACA Medicaid expansion on domestic violence to be reflection of economic independence. That is, expanding insurance program can lower domestic violence through reducing women's economic reliance on spouses.

4.7 Conclusion

This study provides causal estimates of the impact of health care on domestic violence. I find that improved access to health care significantly decreases domestic violence. The results for different measures of accessing health care are consistent, speaking to the strength of the results. Importantly, I exploit variation in women's health insurance exogenously due to differences in insurance premium tax across states, confirming the baseline results. Furthermore, I examine the potential mechanism for why expanding Medicaid may affect domestic violence. My results are consistent with the economic independence account in which the Medicaid expansion decreases women's dependency on their spouses' health insurance plans.

Taken together, this study has several important implications. First, while previous academic research has focused on the causes of domestic violence, here I focus on the delivery of supportive healthcare services to abused women that can contribute to prevention of abuse. Second, these findings suggest that in addition to improved access to health care services, policies that serve to address health inequality also reduce domestic abuse and the costs associated with it. Third, millions of people across the US have become uninsured due to the COVID-19 pandemic. Increasing flexibilities for Medicaid could be potentially important in response to the surge in domestic violence, particularly in states not expanding Medicaid under the ACA.

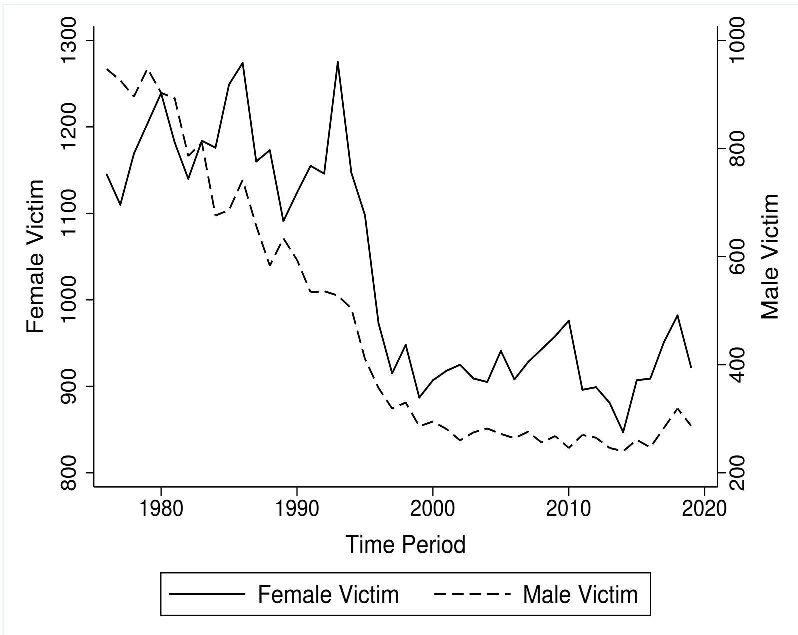


Fig. 1: Trends in Domestic Homicides

Notes: This figure shows the total counts of domestic homicides by victim sex over the sample period from 1976 to 2019 in the US. The left vertical axis measures the number of domestic homicides involving female victims, while the right axis records the number of domestic homicides involving male victims.

Table 1: State Medicaid Expansion

State	Expansion Date
Alaska	9/1/2015
Arizona	1/1/2014
Arkansas	1/1/2014
California	1/1/2014
Colorado	1/1/2014
Connecticut	1/1/2014
Delaware	1/1/2014
District of Columbia	1/1/2014
Hawaii	1/1/2014
Illinois	1/1/2014
Indiana	2/1/2015
Iowa	1/1/2014
Kentucky	1/1/2014
Louisiana	7/1/2016
Maine	1/1/2019
Maryland	1/1/2014
Massachusetts	1/1/2014
Michigan	4/1/2014
Minnesota	1/1/2014
Montana	1/1/2016
Nevada	1/1/2014
New Hampshire	8/15/2014
New Jersey	1/1/2014
New Mexico	1/1/2014
New York	1/1/2014
North Dakota	1/1/2014
Ohio	1/1/2014
Oregon	1/1/2014
Pennsylvania	1/1/2015
Rhode Island	1/1/2014
Vermont	1/1/2014
Virginia	1/1/2019
Washington	1/1/2014
West Virginia	1/1/2014

Notes: This table presents the names of states that adopted the Medicaid expansions through the ACA, and the corresponding dates.

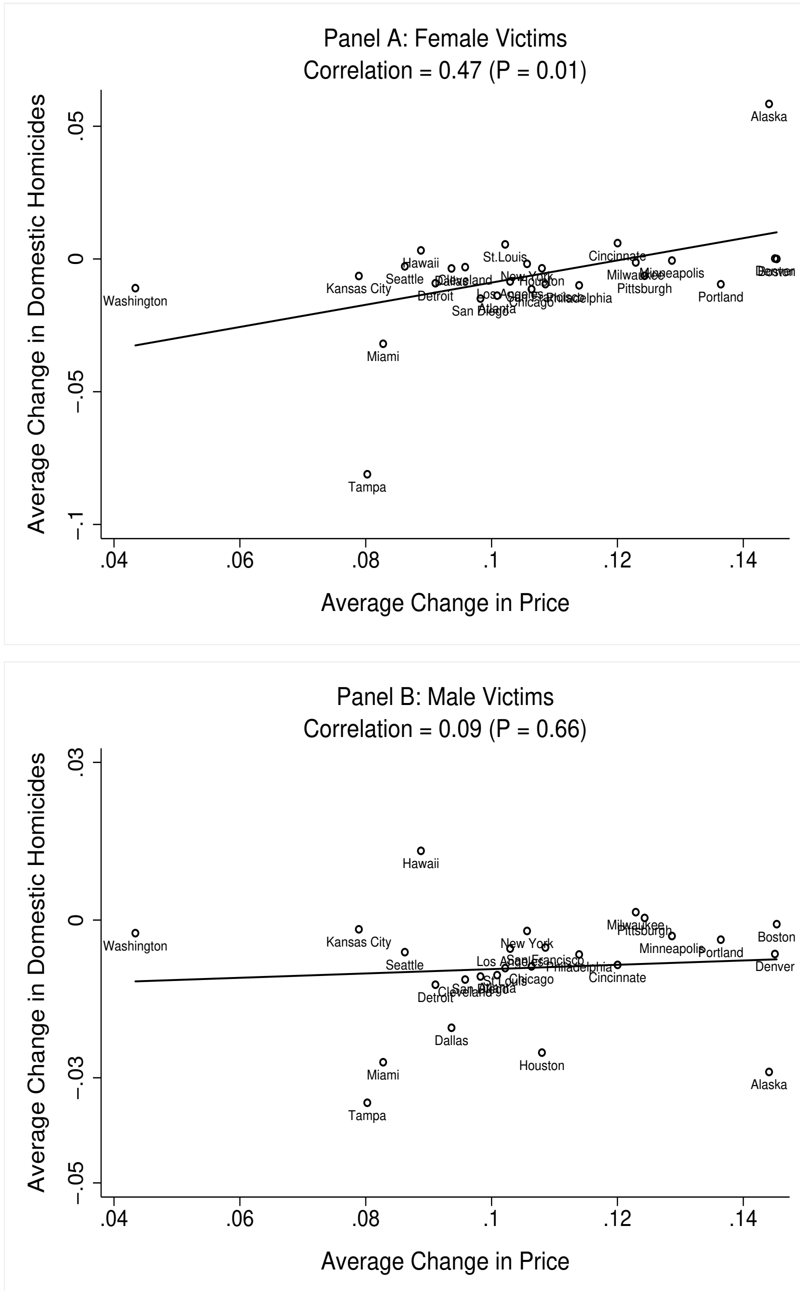


Fig. 2: Domestic Homicide and Medical Care Price

Notes: The figure shows the correlation between the average annual change over the sample period in domestic homicide rate and medical care price at the MSA level. Panels A displays domestic homicides involving female victims against medical care price. Panels B displays domestic homicides involving male victims against medical care price.

Table 2: Summary Statistics of Key Variables

	(1)	(2)	(3)
	N	Mean	S.D.
Panel A: MSA-Year Level			
Domestic Homicides against Women	934	0.384	0.220
(per 100,000 population)			
Domestic Homicides against Men	934	0.174	0.203
(per 100,000 population)			
Medical Care Price	934	2.602	1.333
Panel B: State Level			
Domestic Homicides against Women	50	0.353	0.174
(per 100,000 population)			
Domestic Homicides against Men	50	0.101	0.069
(per 100,000 population)			
Women's Private Insurance Coverage	50	0.670	0.061
Health Insurance Premium Tax (%)	48	2.030	0.675

Notes: Panel A displays summary statistics for annual domestic homicide rates and medical care price between 1976 and 2019 at the MSA level. Panel B presents summary statistics for average domestic homicide rates, average women's private insurance coverage, and premium tax between 2017 and 2019 at the state level. For each variable, I report the observation, the mean, and the standard deviation.

Table 3: Summary Statistics

	(1)	(2)	(3)	(4)
	Expansion States		Nonexpansion States	
	Mean	SD	Mean	SD
Panel A: Domestic Violence				
Domestic Homicides against Women	3.33	3.61	3.83	3.44
Adult Women	3.28	3.54	3.76	3.37
Non-adult Women	0.05	0.24	0.07	0.32
Extensive Margin	0.81	0.39	0.85	0.35
Non-domestic Homicides against Women	7.00	6.50	7.90	6.57
Domestic Homicides against Men	0.76	1.04	1.15	1.28
Panel B: Other Outcome Variables				
Women's Dependent Coverage	0.39	0.05	0.34	0.05
Mental Illness	0.18	0.01	0.18	0.02
Serious Mental Illness	0.04	0.01	0.04	0.01
Panel C: Control Variables				
Police Officers (per 100,000 Persons)	233.05	58.83	248.54	57.17
Female Unemployment rate	0.08	0.02	0.08	0.02
Male Unemployment rate	0.11	0.02	0.09	0.02
Female Income (in $1,000)	28.85	2.87	26.32	1.66
Male Income (in $1,000)	39.75	4.20	36.36	2.53

Notes: The above table presents population weighted means and standard deviations of baseline variables, measured for 2009–2013, in states that expand Medicaid under the ACA and non-expansion states, respectively. Extensive margin represents the monthly probability that at least one domestic homicide against women occurs. Women's dependent coverage is the rate of married women have health insurance coverage, as a dependent, on employment-based health insurance.

Table 4: OLS Estimates of the Effect of Medical Care Price on Domestic Homicide Rate: Female Victims

	Female Domestic Homicide Rate			
	(1)	(2)	(3)	(4)
Medical Care Price	0.076***	0.060***	0.072***	0.070**
	(0.014)	(0.015)	(0.016)	(0.035)
MSA- and Year-Fixed Effects	YES	YES	YES	YES
All Goods Price	NO	YES	YES	YES
Other Controls	NO	NO	YES	YES
MSA-Specific Time Trend	NO	NO	NO	YES
Observations	934	934	934	934

Notes: This table presents OLS estimates of the effect of medical care price on domestic homicides against women per 100,000 population. Other controls include food price, alcohol price, police employment by gender, female officer share, unemployment by gender, personal income by gender, and population. Observations are weighted by MSA population. Significance at the 10%, 5%, and 1% levels is indicated by *, **, and ***, respectively.

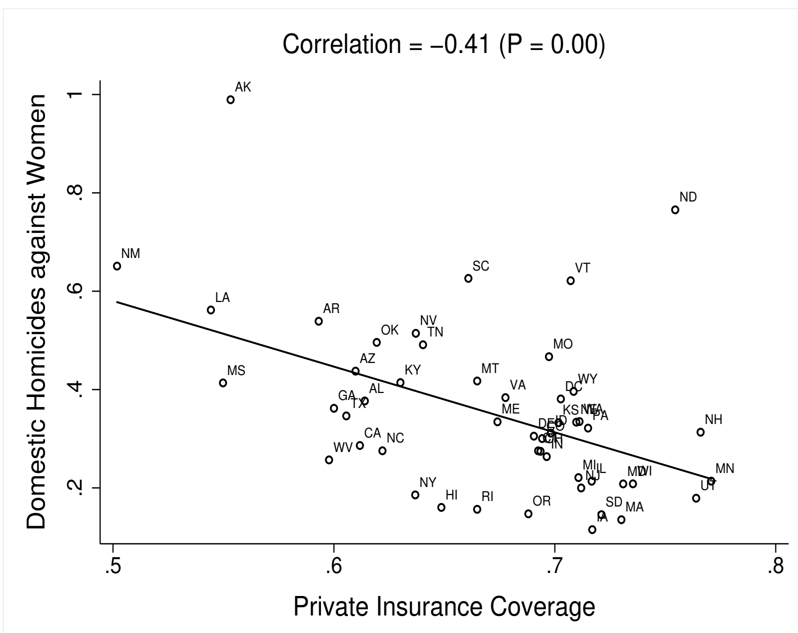


Fig. 3: Female Domestic Homicides and Insurance Coverage

Notes: The figure presents the correlation between average private insurance coverage of women and average female domestic homicide rates between 2017 and 2019.

Table 5: OLS Estimates of the Effect of Medical Care Price on Domestic Homicide Rate: Male Victims

	Male Domestic Homicide Rate			
	(1)	(2)	(3)	(4)
Medical Care Price	0.077***	0.006	0.016	0.011
	(0.011)	(0.012)	(0.013)	(0.024)
MSA- and Year-Fixed Effects	YES	YES	YES	YES
All Goods Price	NO	YES	YES	YES
Other Controls	NO	NO	YES	YES
MSA-Specific Time Trend	NO	NO	NO	YES
Observations	934	934	934	934

Notes: This table presents OLS estimates of the effect of medical care price on domestic homicides against men per 100,000 population. Other controls include food price, alcohol price, police employment by gender, female officer share, unemployment by gender, personal income by gender, and population. Observations are weighted by MSA population. Significance at the 10%, 5%, and 1% levels is indicated by *, **, and ***, respectively.

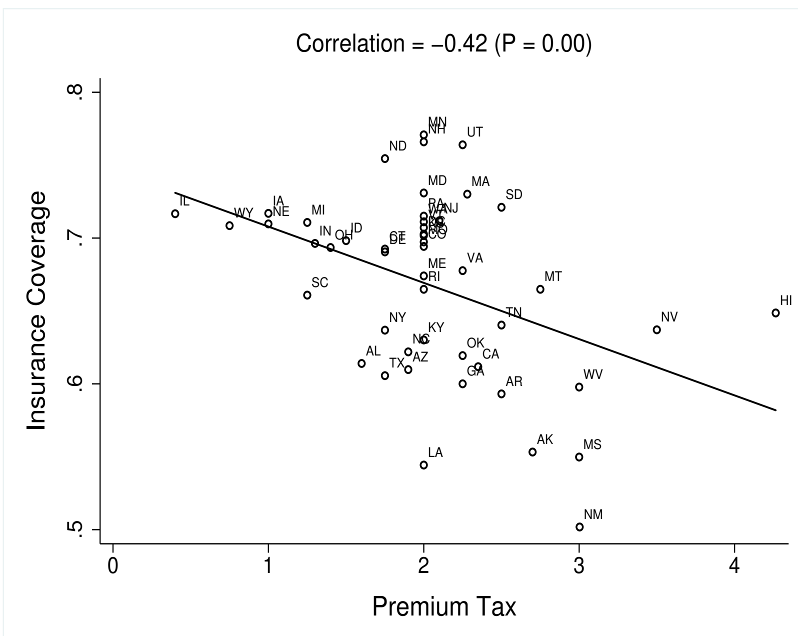


Fig. 4: Insurance Premium Tax and Insurance Coverage

Notes: The figure presents the correlation between insurance premium tax rates and average private insurance coverage of women between 2017 and 2019.

Table 6: OLS Estimates of the Effect of Women's Private Insurance Coverage on Domestic Homicides

	Female Domestic Homicide Rate		
	(1)	(2)	(3)
Private Insurance	-0.985***	-1.309***	-1.473***
	(0.256)	(0.193)	(0.344)
Police Employment by Gender	NO	YES	YES
Other Controls	NO	NO	YES
Observations	50	50	50

Notes: This table presents OLS estimates of the effect of women's private insurance coverage rate on female domestic homicides. Other controls include female officer share, unemployment by gender, personal income by gender, and population. Observations are weighted by state population. Heteroskedasticity robust standard errors are reported below point estimates.Significance at the 10%, 5%, and 1% levels is indicated by *, **, and ***, respectively.

Table 7: Descriptive Statistics of Premium Insurance Tax

	(1)	(2)	(3)
	Mean Value		
	Private Insurance Coverage	Female Domestic Homicide	Number of States
Premium Tax			
Under 1%	0.71	0.26	4
1% to 2%	0.68	0.35	27
2% to 3%	0.65	0.39	12
Over 3%	0.58	0.40	5

Notes: This table presents the average rate of women's private insurance coverage and domestic homicides against women between 2017 and 2019 by ranges of state health premium tax rates.

Table 8: Estimates from the Horse Race Analysis

	Female Domestic Homicide Rate		
	(1)	(2)	(3)
Private Insurance	-0.951***		-0.862***
	(0.263)		(0.288)
Premium Tax		0.056**	0.023
		(0.027)	(0.029)
Controls	NO	NO	NO
Observations	48	48	48

Notes: This table presents estimates of the effect of women's private insurance coverage and health insurance premium tax on female domestic homicides across specifications. Heteroskedasticity robust standard errors are reported below point estimates. Significance at the 10%, 5%, and 1% levels is indicated by *, **, and ***, respectively.

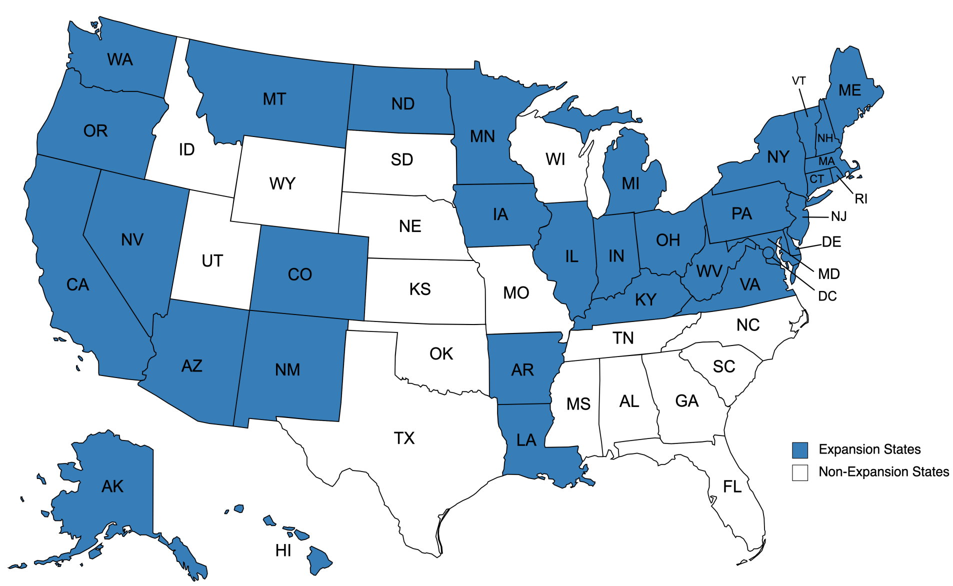


Fig. 5: Status of State Medicaid Expansion Decisions

Table 9: Estimates of the Effect of Insurance Coverage on Female Domestic Homicides

	(1)	(2)	(3)	(4)
	OLS Domestic Homicides	Reduced Form Domestic Homicides	First Stage Private Insurance	IV Structure Domestic Homicides
Private Insurance	-1.350***			-2.068***
	(0.338)			(0.662)
Premium Tax		0.066**	-0.032***	
		(0.027)	(0.011)	
F-Statistic			18.14	
Controls	YES	YES	YES	YES
Observations	48	48	48	48

Notes: This table displays the results from 2SLS, where the dependent variable is the average female domestic homicide rate per 100,000 between 2017 and 2019, the endogenous regressor is the average private insurance coverage for women between 2017 and 2019, and the instrumental variable used is insurance premium tax rate. Controls include police employment by gender, female officer share, unemployment by gender, female officer share, personal income by gender, and population. Significance at the 10%, 5%, and 1% levels is indicated by *, **, and ***, respectively.

Table 10: Estimates of the Effect of Insurance Coverage on Male Domestic Homicides

	(1)	(2)	(3)	(4)
	OLS Domestic Homicides	Reduced Form Domestic Homicides	First Stage Private Insurance	IV Structure Domestic Homicides
Private Insurance	-0.117			-0.274
	(0.124)			(0.349)
Premium Tax		0.009	-0.032***	
		(0.012)	(0.011)	
F-Statistic			18.14	
Controls	YES	YES	YES	YES
Observations	48	48	48	48

Notes: This table displays the results from 2SLS, where the dependent variable is the average male domestic homicide rate per 100,000 population between 2017 and 2019, the endogenous regressor is the average private insurance coverage for women between 2017 and 2019, and the instrumental variable used is insurance premium tax rate. Controls include police employment by gender, female officer share, unemployment by gender, female officer share, personal income by gender, and population. Significance at the 10%, 5%, and 1% levels is indicated by *, **, and ***, respectively.

Table 11: Estimates of the Effect of Insurance Coverage on Homicides against Women

	(1)	(2)	(3)	(4)
	OLS	Reduced Form	First Stage	IV Structure
	Domestic	Domestic	Private	Domestic
	Homicides	Homicides	Insurance	Homicides
Private Insurance	-1.908			-2.438
	(1.303)			(2.169)
Premium Tax		0.030	-0.032***	
		(0.098)	(0.011)	
F-Statistic			18.14	
Controls	YES	YES	YES	YES
Observations	48	48	48	48

Notes: This table displays the results from 2SLS, where the dependent variable is the average rate of homicides against women between 2017 and 2019, the endogenous regressor is the average private insurance coverage for women between 2017 and 2019, and the instrumental variable used is insurance premium tax rate. Controls include police employment by gender, female officer share, unemployment by gender, female officer share, personal income by gender, and population. Significance at the 10%, 5%, and 1% levels is indicated by *, **, and ***, respectively.

Table 12: Estimates of the Effect of ACA Medicaid Expansion on Domestic Homicides against Women

	Female Domestic Homicides			
	(1)	(2)	(3)	(4)
Medicaid Expansion	-0.312*	-0.323*	-0.362*	-0.548**
	(0.180)	(0.182)	(0.197)	(0.250)
State Fixed Effects	YES	YES	YES	YES
Year Fixed Effects	YES	NO	NO	NO
Month Fixed Effects	YES	NO	NO	NO
Year-Month Fixed Effects	NO	YES	YES	YES
Other Controls	NO	NO	YES	YES
State Trends	NO	NO	NO	YES
Observations	6,120	6,120	6,120	6,120

Notes: This table presents DiD estimates of the effect of ACA medicaid expansion on domestic homicides against women. The dependent variable is the number of domestic homicides against women in state s on a given month m. The variable Medicaid Expansion denotes a dummy variable that takes the value one for the post-expansion period. All regressions control for the log of population. Other controls include the log of police rate, unemployment by gender, and income by gender at the state-year level. All SEs are clustered at the state level and observations are weighted by state population. Significance at the 10%, 5%, and 1% levels is indicated by *, **, and ***, respectively.

Table 13: Estimates of the Effect of ACA Medicaid Expansion on Domestic Homicides: Female Victims Aged 18 and above

	Female Domestic Homicides			
	(1)	(2)	(3)	(4)
Medicaid Expansion	-0.294*	-0.306*	-0.343*	-0.519**
	(0.165)	(0.168)	(0.183)	(0.239)
State Fixed Effects	YES	YES	YES	YES
Year Fixed Effects	YES	NO	NO	NO
Month Fixed Effects	YES	NO	NO	NO
Year-Month Fixed Effects	NO	YES	YES	YES
Other Controls	NO	NO	YES	YES
State Trends	NO	NO	NO	YES
Observations	6,120	6,120	6,120	6,120

Notes: This table presents DiD estimates of the effect of ACA medicaid expansion on domestic homicides against women aged 18 and over. The dependent variable is the number of domestic homicides against women in state s on a given month m. All regressions control for the log of population. Other controls include the log of police rate, unemployment by gender, and income by gender at the state-year level. All SEs are clustered at the state level and observations are weighted by state population. Significance at the 10%, 5%, and 1% levels is indicated by *, **, and ***, respectively.

Table 14: Estimates of the Effect of ACA Medicaid Expansion on Domestic Homicides: Female Victims Aged under 18

	Female Domestic Homicides			
	(1)	(2)	(3)	(4)
Medicaid Expansion	-0.018	-0.018	-0.019	-0.029
	(0.024)	(0.025)	(0.023)	(0.019)
State Fixed Effects	YES	YES	YES	YES
Year Fixed Effects	YES	NO	NO	NO
Month Fixed Effects	YES	NO	NO	NO
Year-Month Fixed Effects	NO	YES	YES	YES
Other Controls	NO	NO	YES	YES
State Trends	NO	NO	NO	YES
Observations	6,120	6,120	6,120	6,120

Notes: This table presents DiD estimates of the effect of ACA medicaid expansion on domestic homicides against women. The dependent variable is the number of domestic homicides against women aged under 18 in state s on a given month m. All regressions control for the log of population. Other controls include the log of police rate, unemployment by gender, and income by gender at the state-year level. All SEs are clustered at the state level and observations are weighted by state population. Significance at the 10%, 5%, and 1% levels is indicated by *, **, and ***, respectively.

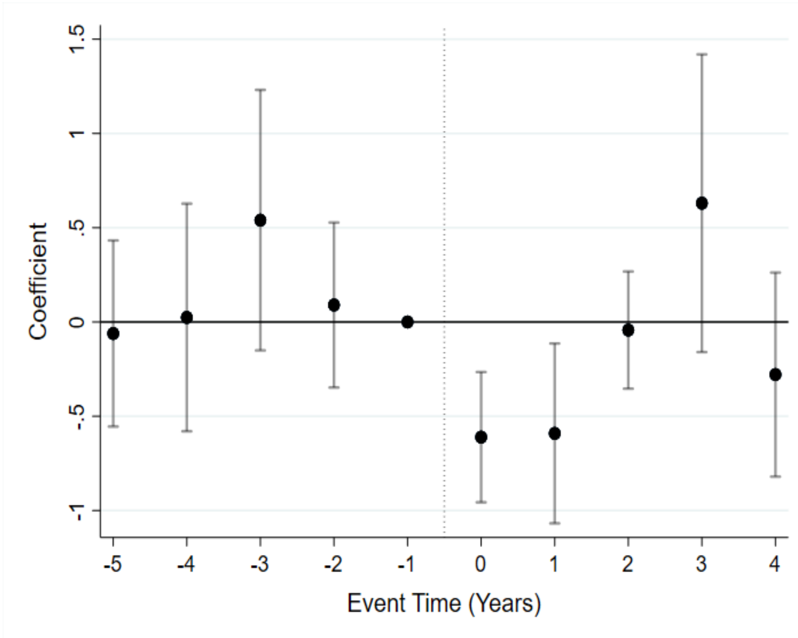


Fig. 6: Event Study Analysis for Domestic Homicides

Notes: The figure presents the event-study plots of the effect of the ACA Medicaid expansion on domestic homicides against women following Sun and Abraham (2021). Standard errors are clustered at the state level and observations are weighted by state population. The dotted vertical line represents the time of treatment. Bands indicate 90 percent confidence intervals.

Table A4.1: Estimates of the Effect of Premium Tax on Public Insurance Coverage

	Medicaid		Medicare	
	(1)	(2)	(3)	(4)
Premium Tax	0.016	0.017	-0.003	-0.003
	(0.013)	(0.013)	(0.006)	(0.005)
Controls	NO	YES	NO	YES
Observations	48	48	48	48

Notes: This table presents estimates of the effect of premium tax on Medicaid and Medicare coverage rates for women. Controls include police employment by gender, female officer share, unemployment by gender, personal income by gender, and population. Heteroskedasticity robust standard errors are reported below point estimates. Significance at the 10%, 5%, and 1% levels is indicated by *, **, and ***, respectively.

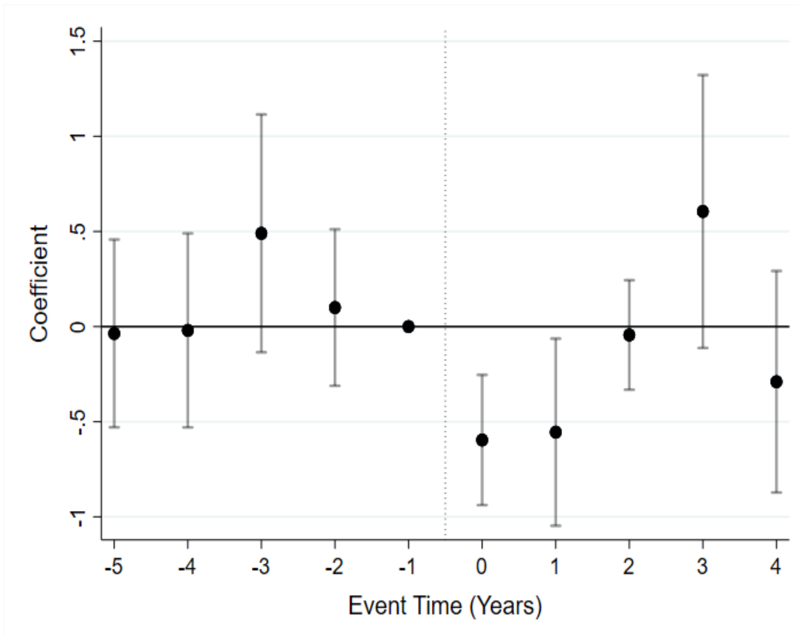


Fig. 7: Event Study Analysis for Domestic Homicides: Female Victims Aged 18 and Over

Notes: The figure presents the event-study plots of the effect of the ACA Medicaid expansion on domestic homicides against women aged 18 and over, following Sun and Abraham (2021). Standard errors are clustered at the state level and observations are weighted by state population. The dotted vertical line represents the time of treatment. Bands indicate 90 percent confidence intervals.

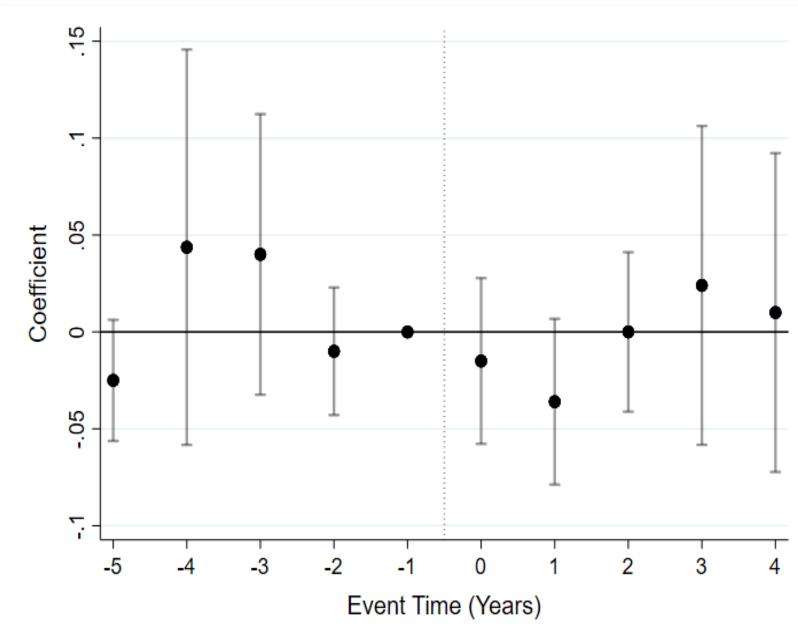


Fig. 8: Event Study Analysis for Domestic Homicides: Female Victims Aged Under 18

Notes: The figure presents the event-study plots of the effect of the ACA Medicaid expansion on domestic homicides against women aged under 18, following Sun and Abraham (2021). Standard errors are clustered at the state level and observations are weighted by state population. The dotted vertical line represents the time of treatment. Bands indicate 90 percent confidence intervals.

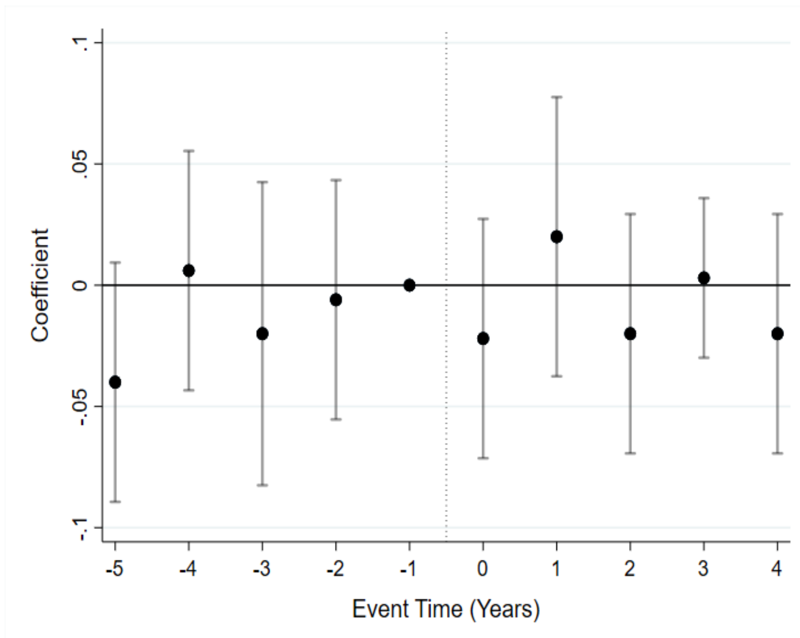


Fig. 9: Event Study Analysis for Domestic Homicides: Extensive Margin

Notes: The figure presents the event-study plots of the effect of the ACA Medicaid expansion on the probability of at least one domestic homicide against women occurring, following Sun and Abraham (2021). Standard errors are clustered at the state level and observations are weighted by state population. The dotted vertical line represents the time of treatment. Bands indicate 90 percent confidence intervals.

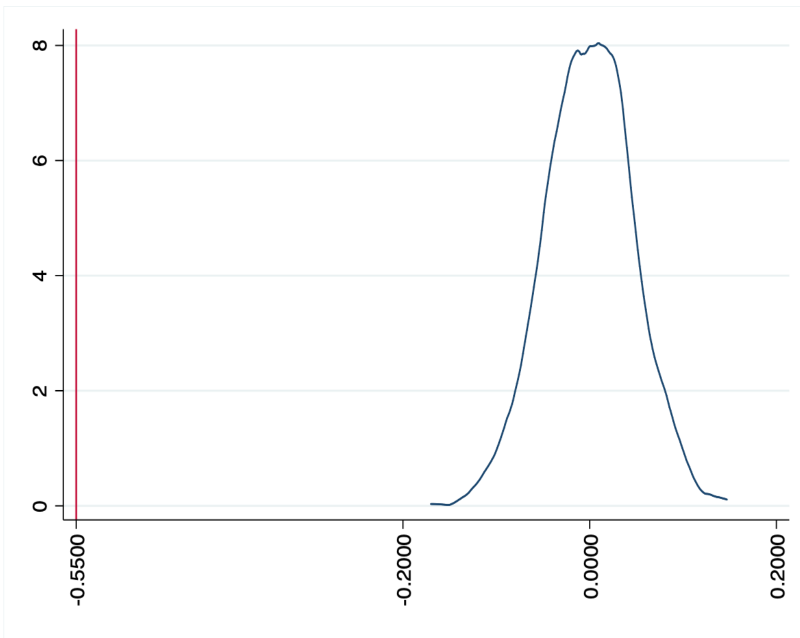


Fig. 10: Distribution of Placebo Estimates

Notes: The figure presents the results of randomising the treatment across states with one thousand permutations. The red vertical line represents the estimated coefficient in my main specification.

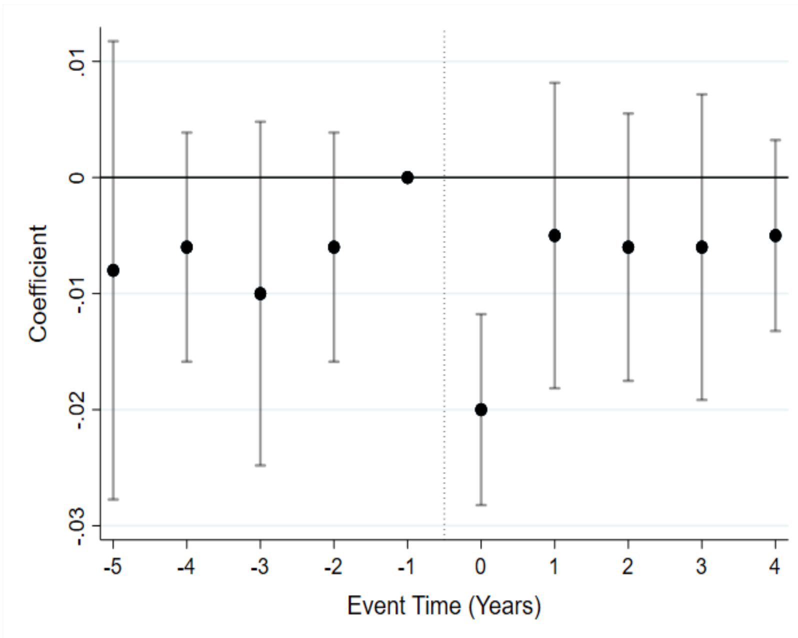


Fig. 11: Event Study Analysis for the Effect of Medicaid Expansion on Women's Dependent Coverage

Notes: The figure presents the event-study plots of the effect of the ACA Medicaid expansion on rates of married women with dependent coverage on employment-based insurance at the state-year level, following Sun and Abraham (2021). Standard errors are clustered at the state level and observations are weighted by state population. The dotted vertical line represents the time of treatment. Bands indicate 90 percent confidence intervals.

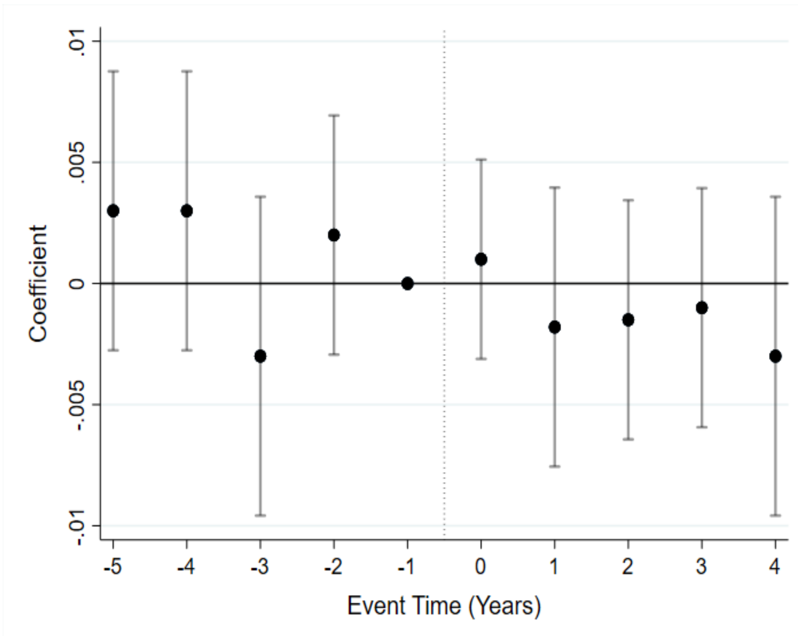


Fig. 12: Event Study Analysis for the Effect of Medicaid Expansion on Female Unemployment

Notes: The figure presents the event-study plots of the effect of the ACA Medicaid expansion on unemployment rates of women at the state-year level, following Sun and Abraham (2021). Standard errors are clustered at the state level and observations are weighted by state population. The dotted vertical line represents the time of treatment. Bands indicate 90 percent confidence intervals.

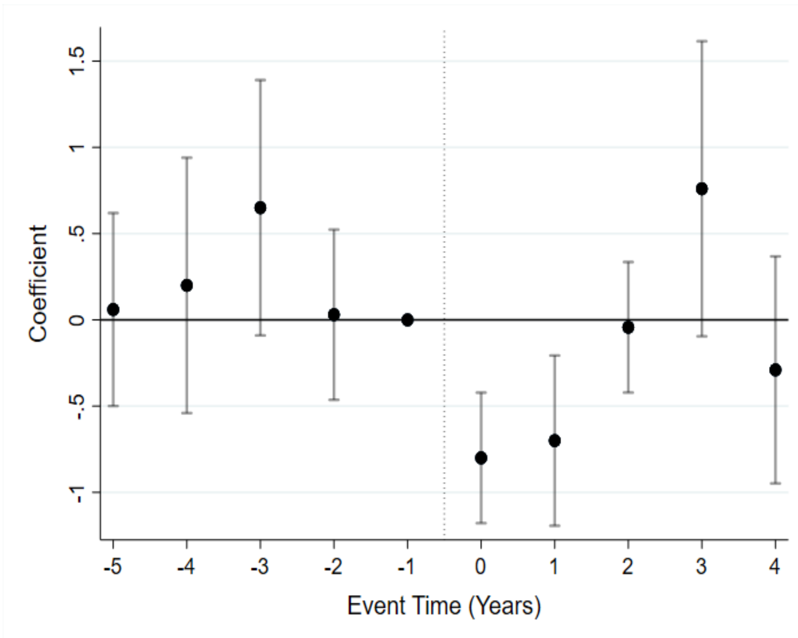


Fig. 13: Event Study Analysis for the Effect of Medicaid Expansion on Domestic Homicides: Above Median Decrease in Dependent Coverage

Notes: The figure presents the event-study plots of the effect of the ACA Medicaid expansion in states with above median decrease in women's dependent coverage on domestic homicides against women, following Sun and Abraham (2021). Standard errors are clustered at the state level and observations are weighted by state population. The dotted vertical line represents the time of treatment. Bands indicate 90 percent confidence intervals.

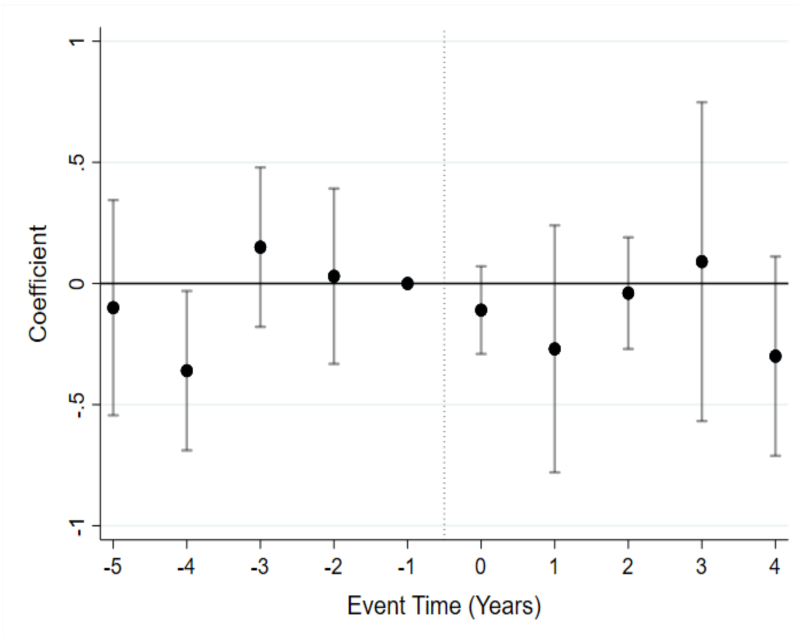


Fig. 14: Event Study Analysis for the Effect of Medicaid Expansion on Domestic Homicides: Below Median Decrease in Dependent Coverage

Notes: The figure presents the event-study plots of the effect of the ACA Medicaid expansion in states with below median decrease in women's dependent coverage on domestic homicides against women, following Sun and Abraham (2021). Standard errors are clustered at the state level and observations are weighted by state population. The dotted vertical line represents the time of treatment. Bands indicate 90 percent confidence intervals.

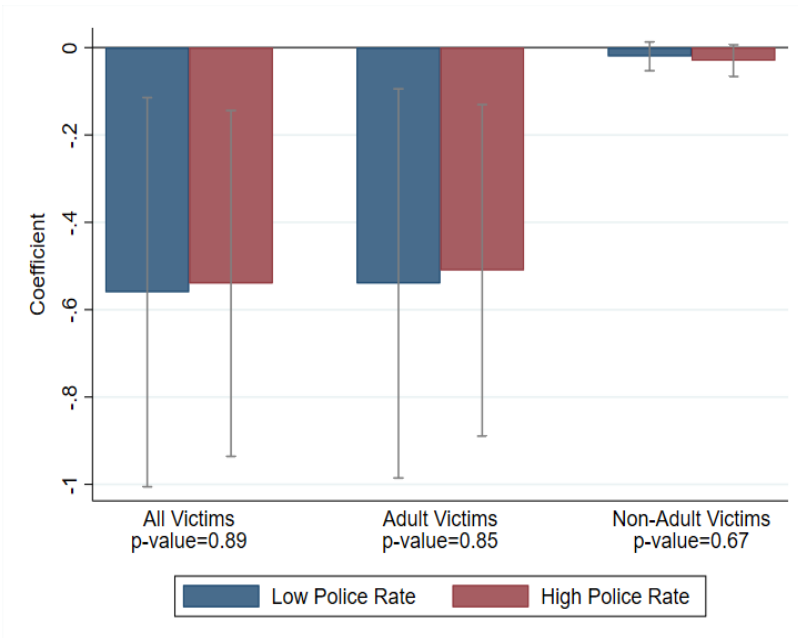


Fig. 15: Heterogeneity by Policing

Notes: The figure presents effects of Medicaid expansion on domestic homicides for all female victims, female victims aged 18 and over, female victims under 18, all at a low level of policing (the twenty-fifth percentile: 189 officers per 100,000 persons) and at a high level of policing (the seventy-fifth percentile: 256 officers per 100,000 persons), respectively. Bands indicate 90 percent confidence intervals. The p-value toward the bottom of the figure is the p-value of a test of the null hypothesis that the low police rate/high police rate effects are equal to one another.

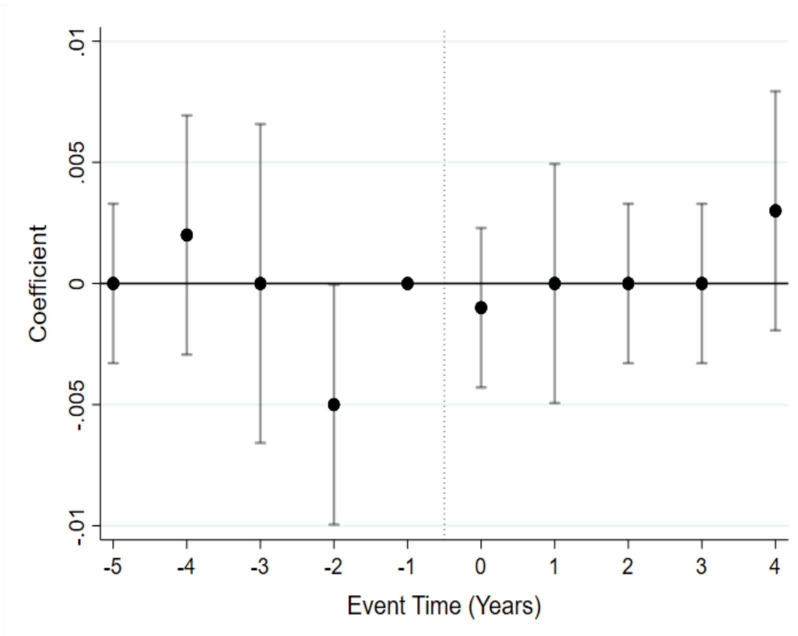


Fig. 16: Event Study Analysis for the Effect of Medicaid Expansion on Prevalence of Mental Illness

Notes: The figure presents the event-study plots of the effect of the ACA Medicaid expansion on the prevalence of mental illness among adults at the state-year level, following Sun and Abraham (2021). Standard errors are clustered at the state level and observations are weighted by state population. The dotted vertical line represents the time of treatment. Bands indicate 90 percent confidence intervals.

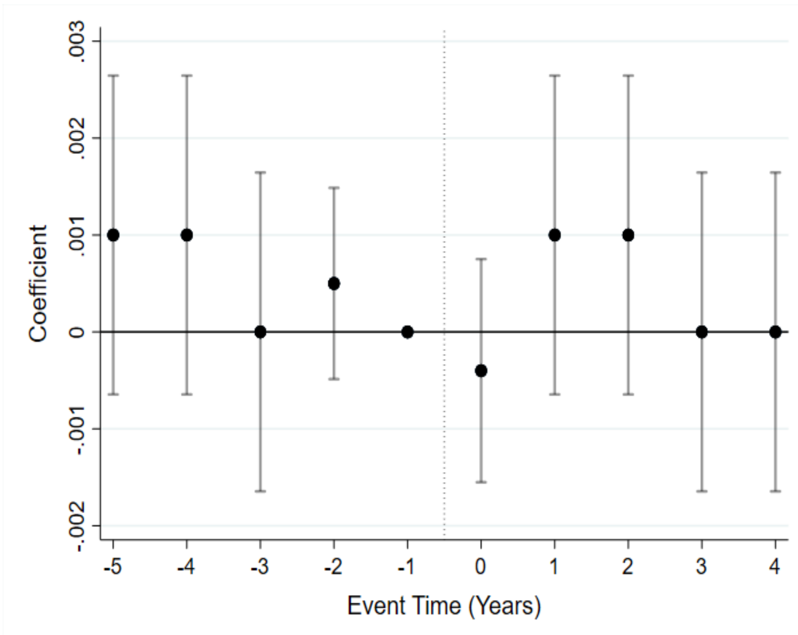


Fig. 17: Event Study Analysis for the Effect of Medicaid Expansion on Prevalence of Serious Mental Illness

Notes: The figure presents the event-study plots of the effect of the ACA Medicaid expansion on the prevalence of serious mental illness among adults at the state-year level, following Sun and Abraham (2021). Standard errors are clustered at the state level and observations are weighted by state population. The dotted vertical line represents the time of treatment. Bands indicate 90 percent confidence intervals.

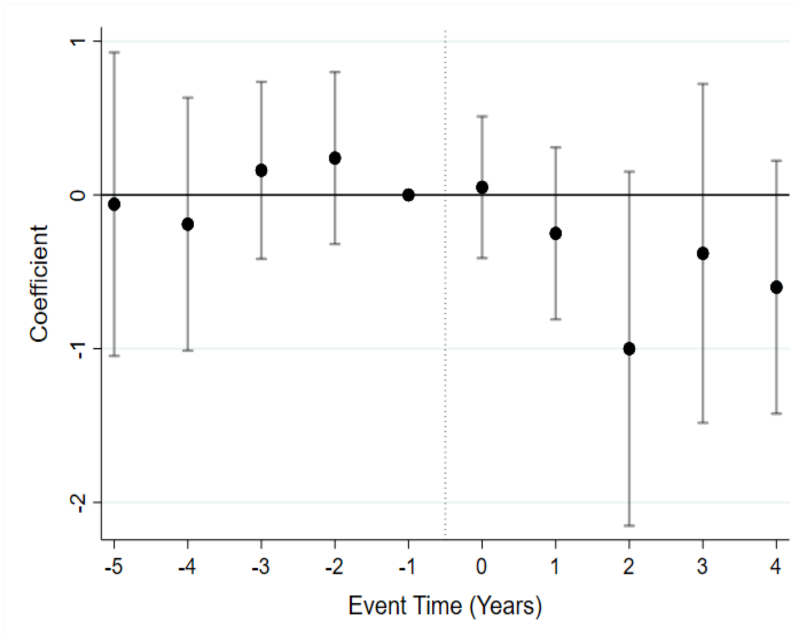


Fig. 18: Event Study Analysis for the Effect of Medicaid Expansion on Non-Domestic Homicides against Women

Notes: The figure presents the event-study plots of the effect of the ACA Medicaid expansion on non-domestic homicides against women at the state-year level, following Sun and Abraham (2021). Standard errors are clustered at the state level and observations are weighted by state population. The dotted vertical line represents the time of treatment. Bands indicate 90 percent confidence intervals.

Table A4.2: Estimates of the Effect of Premium Tax on Women's Economic Welfare

	Female Unemployment		Female Income	
	(1)	(2)	(3)	(4)
Premium Tax	0.002	0.003	507.171	1,375.429
	(0.003)	(0.002)	(1287.493)	(855.187)
Controls	NO	YES	NO	YES
Observations	48	48	48	48

Notes: This table presents estimates of the effect of premium tax on women's labour force participation and women's income. In the first two columns the dependent variable is female unemployment rate, while in the last two columns the dependent variable is mean income of women. The same set of controls are used as the main specification with the outcome variable omitted from them. Heteroskedasticity robust standard errors are reported below point estimates. Significance at the 10%, 5%, and 1% levels is indicated by *, **, and ***, respectively.

Table A4.3: Distribution of Primary Source of Payment for Medical Care Resulting from Domestic Violence against Women

	(1)	(2)
	Rape Victims (Percent Paid)	Physical Assault Victims (Percent Paid)
Payer		
Private Insurance	45.8	48.3
Out of Pocket	29.2	28.6
Medicaid	12.5	11.0
Medicare	N/A	3.0
Other Public Sources	10.4	6.1
Free Clinics	2.1	1.8
Some Other Source	N/A	1.2
Total	100.0	100.0

Source: Centers for Disease Control (2003).

Table A4.4: Estimates of the Effect of ACA Medicaid Expansion on Domestic Homicides against Men

	Male Domestic Homicides			
	(1)	(2)	(3)	(4)
Medicaid Expansion	0.023	0.021	0.018	0.095
	(0.059)	(0.060)	(0.063)	(0.140)
State Fixed Effects	YES	YES	YES	YES
Year Fixed Effects	YES	NO	NO	NO
Month Fixed Effects	YES	NO	NO	NO
Year-Month Fixed Effects	NO	YES	YES	YES
Other Controls	NO	NO	YES	YES
State Trends	NO	NO	NO	YES
Observations	6,120	6,120	6,120	6,120

Notes: This table presents DiD estimates of the effect of ACA medicaid expansion on domestic homicides against men. The dependent variable is the number of domestic homicides against men in state s on a given month m. The variable Medicaid Expansion denotes a dummy variable that takes the value one for the post-expansion period. All regressions control for the log of population. Other controls include the log of police rate, unemployment by gender, and income by gender at the state-year level. All SEs are clustered at the state level and observations are weighted by state population. Significance at the 10%, 5%, and 1% levels is indicated by *, **, and ***, respectively.

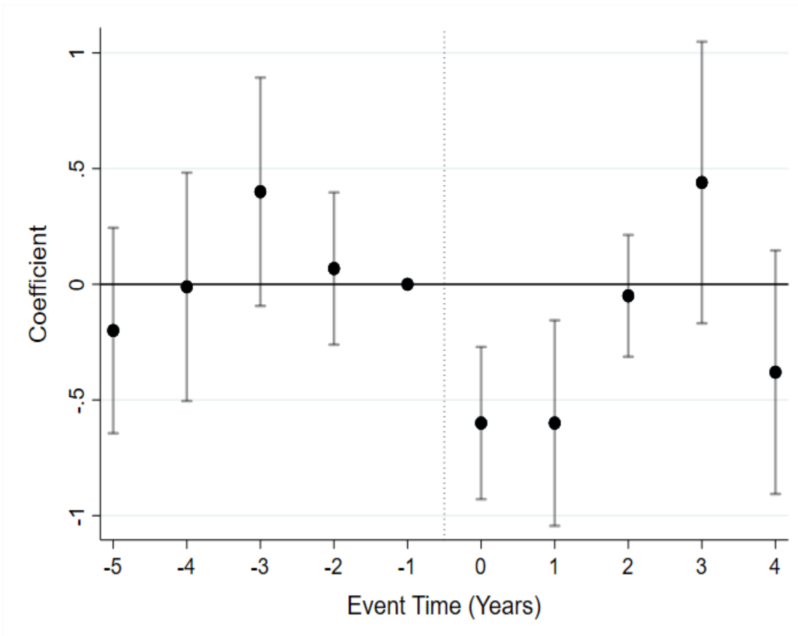


Fig. 19: Event Study Analysis for Domestic Homicides

Notes: The figure presents the event-study plots of the effect of the ACA Medicaid expansion on domestic homicides against women, applying the two-way fixed effects event study method. Standard errors are clustered at the state level and observations are weighted by state population. The dotted vertical line represents the time of treatment. Bands indicate 90 percent confidence intervals.

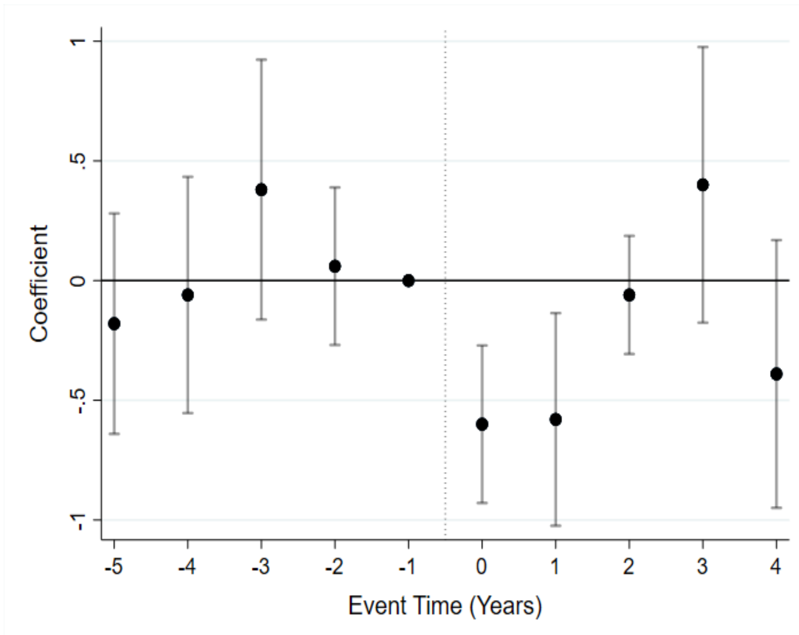


Fig. 20: Event Study Analysis for Domestic Homicides: Female Victims Aged 18 and Over

Notes: The figure presents the event-study plots of the effect of the ACA Medicaid expansion on domestic homicides against women aged 18 and over, applying the two-way fixed effects event study method. Standard errors are clustered at the state level and observations are weighted by state population. The dotted vertical line represents the time of treatment. Bands indicate 90 percent confidence intervals.

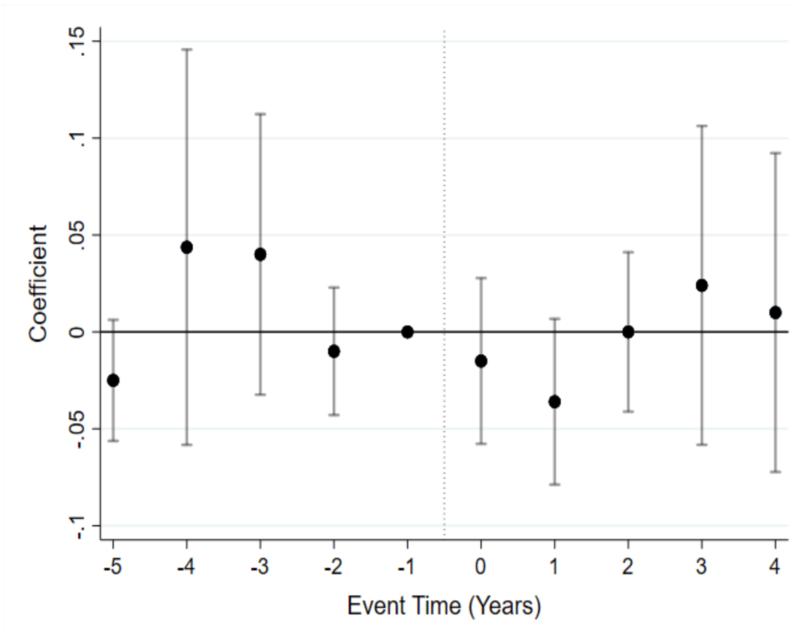


Fig. 21: Event Study Analysis for Domestic Homicides: Female Victims Aged Under 18

Notes: The figure presents the event-study plots of the effect of the ACA Medicaid expansion on domestic homicides against women aged under 18, applying the two-way fixed effects event study method. Standard errors are clustered at the state level and observations are weighted by state population. The dotted vertical line represents the time of treatment. Bands indicate 90 percent confidence intervals.

CHAPTER 5

Punishment and Crime: Evidence from the US Theft Law

Abstract: This study examines the effect of the felony theft thresholds in the U.S. on crime reporting, police performance, and crime incidence. Along these three key dimensions, I find that the felony theft thresholds matter for crime reporting and raising the thresholds weakens the deterrent effect of criminal justice system. Crime victimization data show that offences with theft values above the arbitrary thresholds are more likely to be reported to the police by the victims. These findings survive various robustness checks under different specifications. Furthermore, I find that raising the felony theft thresholds decreases clearance rates of theft and increases theft incidence. These results are consistent with a decrease of deterrence following reduced punishment.

5.1 Introduction

The extent to which criminals respond to changing sanctions is a key pillar of the economic approach to studying criminality. The impact of punishment, as first formally modeled by Becker (1968), reveals a gamble undertaken by a rational criminal weighing up the expected sanction of illegal activity against the benefits. The existing literature has offered support in favor of the crime deterrence effects of the criminal justice system including harsher sanctions and more intensive policing.[1]

However, little is known about the responsiveness of victims to changes in punishment regimes—that is, whether sanctions matter for crime reporting. It is important both in terms of achieving justice for victims and deterring potential crime.[2] In the case of theft (which is the most common crime in the US) around 70 percent of the cases were not reported from 2006 to 2010. Of these, over a third (35 percent) went unreported because the victim believed that the police could not do anything, suggesting gaps in the provision of criminal justice services (Bureau of Justice Statistics, 2012). As Lehman (2022) notes in his article: "stores have learned that their [shoplifting] complaints won't be answered [by the police], so they don't report the crimes."

In this article, I present the first causal evidence on the effect of the severity of punishment on crime reporting and examine police performance and crime incidence outcomes. To accomplish this, I first use a regression discontinuity (RD) design, exploiting the theft law that discontinuously raised the probability of reporting by victims who experience an offence with theft value above an arbitrary threshold. For example, in Missouri, stealing property valued at $750 or more is considered a felony, while larceny of less than $750 is considered a misdemeanor.[3]

The theft felony threshold is attractive from a research perspective due to its simplicity, which causes victims to expect abruptly different police efforts and probability of recovery of losses depending on the theft value. This is conceptually distinct from the question of whether more police efforts can increase deterrence,

[1] These studies have shown how sanction regimes and incarceration can discourage further criminal behavior (Helland and Tabarrok, 2007; Bell et al., 2014; Owens, 2009; Buonanno and Raphael, 2013; Barbarino and Mastrobuoni, 2014; Bhuller et al., 2020) and that an increase in policing lowers crime (Di Tella and Schargrodsky, 2004; Draca et al., 2011; Chalfin and McCrary, 2018; Blesse and Diegmann, 2022). The large literature on punishment and crime is reviewed by Chalfin and McCrary (2017) and Doleac (2020). On the other hand, recent studies have suggested that a criminal conviction has negative labor market effects and leads to an increase in reoffending rates (Mueller-Smith and Schnepel, 2021; Agan et al., 2021; Dobbie et al., 2018; Mueller-Smith, 2015).

[2] Also, law enforcement and community resources may be misallocated due to a lack of accurate information about local crime problems (Bureau of Justice Statistics, 2012).

[3] The theft felony threshold does not apply to recidivism in most states. However, it is not a concern here, as most victims of property crimes do not know the offenders when making decisions on crime reporting (Catalano, 2010).

because perceived victim-specific efforts—the probability of reporting itself deter crime (Goldberg and Nold, 1980).

To conduct the analysis, I use the victimization survey that provides unique data on theft values and crime reporting. An obvious concern is that victims may find difficult in estimating the fair market value of property stolen, which is the premise for theft felony threshold to play a role. It is a probable reason why evidence on the causal effects of the punitiveness of sanctions for thieves is scarce. I overcome this challenge by focusing on cash-only theft in which the face value of stolen cash is generally its market value.

The regression discontinuity estimate indicates a sanction-reporting elasticity: the threshold rule causes thefts with value just above the felony threshold to be around 14 percentage points more likely to be reported. This estimate is economically important, when compared to mean reporting of 14 percentage points in my RD sample. In addition, I conduct a falsification test that the theft felony threshold does not affect reporting of other property crime including robbery and burglary. As such, it is unlikely that my analysis is biased by some unobserved shock contemporaneous with the discontinuous changes in felony eligibility. These results are robust to different controls, alternative bandwidth, and higher polynomials.

Second, leveraging a dynamic difference-in-differences design, I exploit variation in the location and timing of increasing theft felony threshold to compare changes in police performance and crime incidence among states that implement new threshold to states that keep threshold unchanged. To estimate these effects, I employ data on crime clearance rates[4] and recorded crime counts. I find that the clearance rates of theft offences decrease significant by 2 percent one year after raising the theft threshold, and the effect persists for seven years. This result is consistent with the finding for crime reporting presented above, showing the important influence of the sanction regime on police action.

I find considerable evidence that a change in theft threshold policy affects theft incidence. Specifically, the introduction of higher felony theft thresholds brings

[4]Clearance rates have been used as proxy for police performance (McCrary, 2007).

about a mean increase in theft counts by 7 percent per year. In other words, this outcome suggests a significant elasticity of crime with respect to the felony threshold value of 0.05, given that the value of threshold grows by 135 percent on average over the period. These results likely understate the effect, because underreporting of crime could be reinforced or amplified after downgrading felonies. Therefore, my estimates provide important insight into the negative consequences of reducing sanctions on public safety.

This paper closely relates to studies examining the impact of felony theft thresholds on crime. While each of them uses data from the United States, the findings are mixed.[5] Jackson (2020) presents results for a sample of 22 states suggesting that raising thresholds has no impact in thefts in the short run, while it decreases thefts in the long run. Barati (2019) and Bartos and Kubrin (2018) conduct case studies of Arkansas and California, respectively. The former finds that increasing thresholds results in more crime, whereas the latter reports no detectable effects of changing thresholds.

There are two possible reasons why no consensus has emerged. First, these studies draw inferences without considering changes in crime reporting. As mentioned before, lesser punishments could lead to reduced crime reporting. Therefore, ignoring the effect of the harshness of the punishment on crime reporting is making it difficult to interpret the results. In Jackson (2020), for instance, the estimated effect may be interpreted as reflecting a decline in crime incidents reported to police. Second, the estimation samples in case studies are small, which potentially causes concerns about the generalizability of the findings.

In broader term, this paper contributes to the economics of crime literature on punishment and economic incentives. Much of this literature considers the effects of punishment severity on criminal behavior or labor market outcomes (Chalfin and McCrary, 2017; Doleac, 2020; Bhuller et al., 2020). Additionally, there are a handful of studies focusing on the cash flow or return generated by a criminal

[5] There is also one report showing no correlations between changing felony theft thresholds and crime Trusts (2017).

project and finding that if the value of loot from crime rises, thus enhancing the crime return, then the returns from crime go up, which raises crime on the margin (Draca et al., 2019; Draca and Machin, 2015).

Finally, this article builds on a large literature across the social sciences exploring causes of underreporting of crime. The factors include the cost of loss (Clarke and Harris, 1992), the probability of recovery of loss (Goldberg and Nold, 1980), lack of confidence in criminal justice system (Boateng, 2018), unemployment (MacDonald, 2001), insurance coverage (MacDonald, 2002), cultural beliefs (Njuki et al., 2012), and fear of retaliation (World Health Organization, 2005; Kishor and Johnson, 2005).

The remainder of the paper proceeds as follows. Section 2 provides a description of the felony theft thresholds. Section 3 presents the hypothesized effects. Section 4 describes the data construction. Section 5 conducts the analysis of the felony theft thresholds and crime reporting. Section 6 presents the analysis of the effect of raising such thresholds on police performance and crime incidence. Section 7 concludes.

5.2 Background

The basics of the U.S. larceny law, including the distinction of larceny from other types of theft and the division between felony and misdemeanor larceny, were developed before the American Revolution (Rice, 2017). In the face of substantial prison costs, particularly for the past decades, many states have started to raise their thresholds for felony theft, which is a dividing line between felony and misdemeanor theft based on the theft value. For example, in Alabama, the requisite amount for enhancement to a felony was $250, while this amount was changed to $500 in 2003. There is a general provision relating to the valuation of property that defines value as market value or the cost to the victim of replacing the property.

However, these thresholds are not to be applied on all theft incidents. Some kinds of property are usually protected by felony theft penalties without reference to the value of the property stolen such as automobiles and other motor vehicles, airplanes, and firearms. Moreover, since the person who commits misdemeanor theft following a prior felony or theft has shown sufficient evidence of criminality, sentencing recidivists does not comply with the threshold rules at all (Culver, 1969).

Felonies commonly differs from misdemeanors by the length of prison sentence.[6] Bureau of Justice Statistics (1987) provides definitions for them: a felony is a crime punishable by death, by imprisonment for life, or by imprisonment for more than 12 months, while a misdemeanor is a crime punishable by imprisonment for a maximum period of 12 months. Actually, for misdemeanor incidents, prosecutors have a great degree of flexibility in deciding how to punish offenders, and what kinds of plea bargains to negotiate.

Misdemeanors are typically dismissed as low-level events that do not deserve the attention or due process accorded to felonies (Natapoff, 2011). For instance, according to Agan et al. (2021), 21% of nonviolent misdemeanor cases are not prosecuted; the remaining 79% are prosecuted. 74% of nonviolent misdemeanor cases that are prosecuted are eventually disposed of without criminal convictions in Suffolk county.

5.3 Hypothesized Effects

I present my main hypotheses for the effects of felony theft thresholds on crime based on Becker (1968) and Goldberg and Nold (1980) that considers the behavior of three agents relevant to a crime situation: criminals, victims, and police, as shown in Figure 1. Given that changes in behavior of the police would have impact on behavior of both victims and criminals, I focus on the potential effects of felony thresholds on police behavior at first.

[6]Some states also specify imprisonment in the state prison for felonies (Bureau of Justice Statistics, 1987).

5.3.1 Police

I simply view that the goal of the police service is to maintain order and save property by bringing criminals to justice and preventing crime. As in Benson et al. (1998), due to competing demands for scarce police resources, police have to decide how best to allocate the resources across crime cases to achieve the goal. In this calculus, police would invest more in cases that they value higher and consider more serious. As a consequence, the probability of bringing criminals in justice and the recovery of losses for a given theft would depend on its severity. Lee and McCrary (2017) also highlight that it is plausible that law enforcement may exercise discretion in formally investigating a case, based on its severity.

In what way do felony theft thresholds affect police action? The legal system devalues misdemeanor convictions.[7] Police typically treat misdemeanors as unimportant relative to felonies, and misdemeanors often do not trigger jail time (Natapoff, 2015). If a state raises the felony threshold, which would result in many felony-eligible thefts being reduced to misdemeanors, then more thefts could be taken less seriously by police and the overall clearance rate of thefts may decrease.

5.3.2 Victims

One the one hand, victims would benefit from reporting if offenders are punished and losses of victims are recovered. On the other hand, reporting involves opportunity time costs, which may be appreciable, as reporting often has to be done in person and in some cases may involve police visits.

In what way do felony theft thresholds affect crime reporting? As presented above, police are more likely to investigate the felony cases and recover losses for felony theft victims. Accordingly, for felonies, the benefit from reporting would be higher and more reporting would be expected. In the same vein, a marginal decline in punishment severity may induce police to be less responsive to more theft incidents, further discourage reporting. In addition, reduced sanctions may

[7] It does so explicitly by labeling them "petty" and implicitly by withholding public resources from the vast misdemeanor process (Natapoff, 2011).

have unfavourable impact in people's confidence in criminal justice institutions, which could lead to greater underreporting of crimes regardless of the economic benefits or costs from reporting (Boateng, 2018).

5.3.3 Criminals

Criminals make the decision whether to commit a crime by weighing up the expected sanction against the benefits. Thefts yield payoff, which is the value of stolen property in case of thefts. There would be costs that are imposed on thieves, if they are arrested and brought in justice.

In what way do felony theft thresholds affect the behavior of criminals? As described above, in comparison to misdemeanor thefts, felony thefts are associated with more reporting and greater police responsiveness, while they yield higher returns. In this regard, the directional effect of felony theft thresholds on crime is ambiguous. Nevertheless, because raising felony theft thresholds would decrease the responsiveness of police and drive reporting down, potential criminals would only see an increase in the expected benefits of their crime, and tend to commit a crime.

Overall, felony theft thresholds could affect police action and crime reporting via signalling the severity of the crime. While conversely, it is an open question whether or not discontinuities exist across the felony thresholds. In terms of the shift of the sanctions discontinuity, a decline in law and order would be expected due to reduced deterrence following an increase in the an arbitrary amount of the felony theft threshold.

5.4 Data

5.4.1 Crime Victimization Survey

Data on crime reporting are obtained from the Bureau of Justice Statistics' (BJS) National Crime Victimization Survey (NCVS), which is the nation's primary source of information on criminal victimization. I use the extract file created from the NCVS for a sample of Metropolitan Statistical Areas (MSAs) from 1970s through 2000s.[8] This survey contains select household, person, and incident variables for crime victims. In case of theft, it includes information on stolen item(s) and the value of stolen item(s). Considering the potential concern about misvaluation of item(s) by victims, my main analysis focuses on cash-only thefts.

The outcome, crime reporting, is measured by an indicator variable of whether or not the crime was reported to the police. I also extract information on victim and incident characteristics such as gender, race, Hispanic origin, education attainment, household income, incident places, knowledge of offenders, and number of offenders.

5.4.2 Uniform Crime Reports

The data source for the state-year police performance and official records of thefts is the FBI's Uniform Crime Reporting (UCR) System, which is the nation's primary source for them. Following McCrary (2007), I use clearance rates that are defined as the number of crimes cleared by arrest over total crimes, as proxy for police performance.[9] Theft measures reflect the total count of thefts known to police to have occurred during the calendar year.

The UCR contains data for theft exclusive of motor vehicle theft, and motor vehicle theft, separately. Given that motor vehicle thefts are felonies regardless of the value, thefts represent theft offences exclusive of motor vehicle theft in my analysis. To conduct the placebo test, I also collect data on clearance and crime measures for other types of crime excluding theft. In doing so, I am able to examine whether changing felony theft threshold affects thefts exclusively.

[8]Since there are several MSAs covering counties in different states, I match them to state felony threshold data based on the core county with the largest population. My results are robust to exclusion of these MSAs.

[9]In the UCR Program, a law enforcement agency reports that an offense is cleared by arrest, when three specific conditions have been met: (a)arrested; (b)charged with the commission of the offense; (c)turned over to the court for prosecution (whether following arrest, court summons, or police notice) (Federal Bureau of Investigation, 2020).

5.4.3 Felony Theft Thresholds

I obtain data on the value of felony theft thresholds between 1979 to 2004 from the state legislature archives to match crime victimization data.[10] The corresponding thresholds for 19 states are used to match the NCVS sample. Although each state specifies a dollar figure for the division of felony and misdemeanor thefts that is an arbitrary amount, the felony threshold may not be the same value as the laws require that the value of a felony theft must exceed such a figure in some states. For example, a theft of property of the value of more than $200 is a felony in New Jersey. For those states, I use the dollar figures plus 0.0001 as the felony theft thresholds.[11] Figure 2 shows the distribution of the thresholds in 1979 and 2004, respectively.

To conduct analysis on crime clearance and records, I collect additional data on the timing of raising felony theft thresholds for 51 states from Trusts (2017), which are available from 2001 to 2016 and matched to the UCR data.

I supplement my analysis with additional data on police employment information come from the Law Enforcement Officers Killed or Assaulted (LEOKA) collection. To take account of other social-economic characteristics, I also collect data on population, race, education, income per capita and unemployment rate from the U.S. Census.

5.5 Analysis of Crime Reporting

I obtain causal identification from the law by which a theft offence is eligible for a felony conviction. As previously discussed, a theft is eligible for felony with the theft value greater than the felony threshold. While crime reporting may be influenced by characteristics of victims and incidents, these factors do not change discontinuously at the thresholds. As long as the eligibility rule is considered to any degree, the likelihood of treatment will discontinuously increase at the threshold,

[10] There are 21 states in the NCVS sample. As information on thresholds is not available, Oregon and District of Columbia are excluded in the analysis.

[11] In the case of New Jersey, the threshold is $200.0001. The results are robust to using 0.1, 0.01, or 0.001.

making it possible to estimate the effect of the severity of punishment on crime reporting using a fuzzy regression discontinuity design. I pool victims of theft according to the felony thresholds that are applied in each state, so the running variable is the theft value minus the treatment threshold. Victims could find it difficult to evaluate the fair market value of stolen property, so I limit the sample to cash-only thefts.

5.5.1 Regression Discontinuity Specification

Before examining the plausibility of the RD identifying assumptions in detail, I specify the regression model. If I have data for actual felony sentence ideally, I am able to estimate the effect of felony sentence by using a fuzzy RD design. The basic idea behind the RD method is that the felony sentence is determined at least partly by the value of a forcing variable, which is theft value, being on either side of a fixed felony threshold. Nevertheless, the probability of felony sentence does not change from zero to one at the felony threshold, because factors other than the threshold rule affect the felony sentence, such as criminal records and stringency of judge. As in Lee and Lemieux (2010), since there is imperfect compliance of the threshold rule, the RD gap can no longer be interpreted as an average treatment effect (ATE).

Therefore, if I have data for actual felony sentence ideally, the estimated effect is a local average treatment effect (LATE), that is, it is the effect for those thefts with the theft value reaching the felony threshold. Three equations are presented to explain the interpretation:

$$Reporting_{ist} = \alpha_1 Felony_{ist} + g(Value_{ist}, \mu) + \alpha_2 X_{ist} + u_s + \rho_t + \varepsilon_{ist} \tag{1}$$

$$Felony_{ist} = \theta_1 Cutoff_{ist} + f(Value_{ist}, \lambda) + \theta_2 X_{ist} + \eta_s + q_t + \nu_{ist} \tag{2}$$

$$Reporting_{ist} = \beta_1 Cutoff_{ist} + h(Value_{ist}, \sigma) + \beta_2 X_{ist} + w_s + \iota_t + \upsilon_{ist} \tag{3}$$

Where in equation (1), $Reporting_{ist}$ is an indicator equal to 1 if crime i is reported to police in state s in year t. $Felony_{ist}$ is an indicator denoting felony sentence. $Value_{ist}$ represents the value of cash stolen and $g(Value_{ist}, \mu)$ is a polynomial in theft value. X_{ist} is a set of controls for victim and incident characteristics, which is not necessary for identification but improve the efficiency of the estimation. u_s and ρ_t are state and year fixed effects, respectively, and ε_{ist} is the error term. α_1 estimates the effect of felony on crime reporting.

Where in equation (2), $Cutoff_{ist}$ is an indicator of whether the value of cash stolen exceeds the felony threshold, and $f(Value_{ist}, \lambda)$ is a theft value polynomial. θ_1 estimates the effect of being above felony threshold on felony sentence. Combining equation (1) and (2) gives the reduced form equation (3). Because I do not have data for actual felony sentence, $Felony_{ist}$, I cannot estimate α_1 in the equation (1), which is the effect of felony sentence. I am, on the other hand, estimating β_1 in the equation (3), that is the effect of being above the felony theft value threshold by using a sharp RD design. Thus, my sharp RD estimates can be interpreted as an average treatment effect weighted by the probability of being close to the felony theft value threshold.

Following Imbens and Lemieux (2008) and Gelman and Imbens (2019), my primary specification uses local linear regression within a given bandwidth of the treatment threshold, and controls for the running variable (the amount of stolen cash) on either side of the threshold. The optimal bandwidth according to the method of Calonico et al. (2014) is \$304. Results are strikingly similar with different controls, a rectangular kernel, or alternate bandwidths.[12]

Identification requires that relevant factors besides treatment vary smoothly at the threshold between a felony and misdemeanor. This assumption is needed for thefts with cash values just below the threshold to be appropriate counterfactual for those with cash values just above the threshold. To assess the plausibility of the identifying assumptions, Table 1 examines whether victim and incident characteristics are balanced across the felony theft threshold. Columns 1 reports the regression discontinuity estimates, following the main estimating equation, of the

[12]Gelman and Imbens (2019) suggest that high-order polynomials should not be used in RD design, I report the main results by using linear RD polynomials. Nevertheless, my results shown in Appendix Table A5.1 are robust to higher polynomials.

effect of being above the felony threshold on these characteristics. Column 2 shows the clustered standard errors at the state level in parentheses. Column 3 is the p-value for a RD estimate. I find that the estimates are statistically identical to zero, suggesting that victim and incident characteristics are continuous at the treatment threshold. Also I present the graphical version of the balance test and Appendix Figures A5.7.1 to 6 shows analogous RD plots for each characteristic.

Identification also requires the absence of selective sorting around the felony and misdemeanor threshold. This assumption would be violated, for example, if thieves choose to bunch just below the cutoff that would later affect crime reporting. To formally test for sorting, I implement the McCrary test McCrary (2008). Figure 3 shows that the density of the stolen cash distribution is continuous across the threshold.

The absence of selective sorting is also plausible, which can be explained by two factors. The first is that the availability of property is a key determinant of the values of stolen property. The extensive literature has shown that the increase in the availability of items leads to an increase in the theft rate of those items (Copes, 1999; Mansfield et al., 1974; Gould, 1969). The second factor builds on the importance of costs of adjusting to the cutoff. The criminals could face high adjustment costs when bunching below the cutoff. Specifically, extra time and efforts are necessary for a theft to steal the property with the selected value, which thereby could draw victim's attention to the ongoing incident and increase the risk of apprehension. This is analogous to the findings in Jacobs and Cherbonneau (2014) and Copes (1999) that thieves try to minimize the risk by stealing something that was not too noticeable and having a convenient way to remove the object.[13]

[13]This stands in contrast to the finding in Lepage (2020) that offenders respond by bunching to the left of the drug sentencing threshold, as the adjustment costs for offenders carrying the quantities of drug just below the threshold rather than just above the threshold are relatively low.

5.5.2 Main Results

I examine the effect of having the theft value over the felony threshold on crime reporting propensities. As such, this approach answers a fundamental policy question: Does the sanction regime matter for crime reporting? Table 2 presents the estimated effect of stealing cash with the amount over the felony threshold. Column 1 reports the RD estimate without controls and the coefficient is statistically significant at the 1 percent level. It suggests that stealing cash with the amount over the felony threshold increases reporting by 14 percentage points, which is a sizable increase in reporting when considered in absolute terms or relative to average reporting rates (14%). Columns 2 and 3 report results with victim characteristics controls and with all controls, respectively. Through the different specifications, the coefficients remain statistically significant, strongly implying that the estimated effect is thus quite insensitive to the inclusion of control variables.

Figure 4 plots means of reporting rates in bins and predicted reporting rates according to Calonico et al. (2014) highlights the stark changes in crime reporting which occur at the felony threshold. The black lines represent the fitted regressions. There is a notable jump in reporting at the threshold, which is very close to the table point estimates in Table 2. Table 3 presents estimates using the main estimating equation at various bandwidths. The estimates are very robust to different bandwidth choices.

Although the density test shows that the theft value distribution is continuous at the felony threshold, I conduct a "donut" RD test that ignores data close to the threshold to further verifies that my estimates are not driven by nonrandom sorting (Barreca et al., 2011; Bharadwaj et al., 2013; Card and Giuliano, 2014). In doing so, I keep comparing the mean estimates because they approach the theft threshold from each side, while allowing for the possibility that those at the heap are systematically different from surrounding observations. In Table 4 I exclude observations within \$1, \$10, \$50, and \$100 of the felony threshold, respectively. The estimates show that the reporting effect remains in place and are not sensitive to the observations at the threshold.

It is natural to worry about whether rounding up (or down) of the stolen cash value may influence my results. To deal with rounding in this context, I examine the sensitivity of my results when taking account into multiples of \$5 and \$10. Columns 1 to 3 of Table 5 presents the corresponding estimates when controlling for multiples of \$5, multiples of \$10, and both, respectively. The results are quite stable when including additional controls for rounding values.

Furthermore, I perform a placebo exercise where I replace the indicator of theft reporting with that of other property crime reporting. Since burglary and robbery are felonies regardless of the value, I would expect that crossing the felony threshold does not influence reporting of them. Table 6 shows the corresponding results. I do not find any meaningful discontinuities in reporting for other property crime in the placebo exercise.

By and large, victims do respond to the sanction discontinuity and are more likely to report the incident to police when experiencing a more severe crime. The substantial increase in reporting is initial evidence that the increase in punishments and sanctions at the thresholds is effective in strengthening deterrence. Moreover, these results imply that the estimates of the deterrent effect of harsher sanctions in the literature could be a lower bound of the true effect, as harsher sanctions cause more victims to come forward to report the crime.

5.5.3 Placebo Thresholds

In this section, I perform tests for jumps at non-discontinuity points. Following the recommendations of Imbens and Lemieux (2008), these tests are conducted by including the subsamples only to the right or only to the left of the true felony thresholds, so as to avoid including in the estimation a point where there is a discontinuity. First, I test for jumps at the median value of the two subsamples on either side of the felony theft threshold value. If the effect observed at the true felony theft threshold is a genuine one, then there should be no effect observed at other thresholds. Because the felony theft threshold differs by states, the placebo thresholds are presented as relative values to the true felony thresholds. Panel A of Table 7 reports the corresponding results for subsamples of the left and right of the true felony threshold, respectively. As expected, the results do not show any significant coefficients.

Second, I split the sample again at the new median thresholds and create another threshold at the median of the new samples. Panel B of Table 7 shows the corresponding results for those newly generated four thresholds, respectively. I find no evidence for significant discontinuities at these placebo thresholds. Hence, I conclude that discontinuity estimates at the actual felony theft threshold reported are not spurious.

5.6 Analysis of Raising the Felony Theft Thresholds

As discussed above, reduced punishment could decrease police responsiveness and tempt the offenders to commit more crime. Having focused on the relationship between the felony theft thresholds and crime reporting, I now investigate whether changing such thresholds affects police performance and crime incidence. In doing so, my analysis is able to provide important implications for ongoing policy debates over criminal justice reforms. The section first discusses the identification strategy. It then examines whether raising the felony theft thresholds induces a decline in police performance. Finally, it explores the impact on crime incidence. Table 8 presents the years of changing the felony theft thresholds and corresponding state names.[14]

5.6.1 Empirical Strategy

To estimate the effect of higher threshold values on behavior of police and criminals, I intend to exploit the variation in placement and timing of raising the thresh-

[14]Appendix Figure A5.7.7 displays the change in distribution of the threshold value for those states over the time period.

old values.[15] A recent body of research has investigated the biases of the two-way fixed effects difference-in-difference (DD) estimator associated with heterogeneous treatment effects when treatment timing is staggered (Sun and Abraham, 2021; Callaway and Sant'Anna, 2021; Athey and Imbens, 2021; Goodman-Bacon, 2021). In particular, such an estimator equals a weighted average of all possible simple 2x2 DDs that compare one group that changes treatment status to another group that does not. In other words, applying the two-way fixed effects DD framework might lead to a bias in presence of heterogeneous treatment effects across states increasing the felony theft thresholds at different points in time.

As a result, I follow Sun and Abraham (2021), who develop DD estimators of the cohort-and-period specific effects that only rely on the parallel trends assumption, and that are robust to heterogeneous treatment effects.[16] Their estimators categories units into different cohorts based on their initial treatment timing, to avoid the estimates of lags and leads being contaminated by effects from other periods.[17] Specifically, they define their estimates as interaction-weighted (IW) estimates, which are generated mainly in two steps. First, the cohort average treatment effects on the treated (CATT) are computed by estimating the cohort-specific average difference in outcomes relative to never being treated. Second, their estimator estimates a weighted average of CATT with weights equal to the share of each cohort in the relevant period(s).

Based on their methodology, I estimate the following equation:

$$Y_{st} = \sum_{e=1}^{E} \sum_{\substack{k=-10 \\ k\neq -1}}^{10} \beta_{e,k}(\mathbb{1}\{E_s = e\} Post_{st}^{k}) + +\theta X_{st} + \sigma_s + \phi_t + \varepsilon_{st} \tag{4}$$

Where Y_{st} is the outcome of interest in state s and year t. $Post_{st}^{k}$ is a set of relative event-time dummies, that take value of 1 if time t is k years after (or before,

[15]As shown before, there are no states decreasing the felony theft thresholds.

[16]Nevertheless, as a robustness check, I also use the two-way fixed effects (TWFE) DD model. The main results are presented in Appendix Figure A5.7.8, and they are highly similar to the results from using estimators of Sun and Abraham (2021).

[17]Their estimators use the never-treated units are used as controls if there are never-treated units. Their procedure ensures nonnegative weights and better sheds light on dynamic treatment effects.

if k is negative) the increase in the theft felony threshold.[18] e stands for cohorts, different years in which states are treated. Time windows span periods of one year each. Particularly, $\pm k$ ranges from 10 and -10 to respectively 1 and -2 years, as -1 is omitted. Each lag(lead) takes the value of the main regressor -$k(k)$ years away from the increase in the theft felony threshold.

In addition, X_{st} represents a set of controls at the state level, including population, police size, black population share, education, unemployment, and income. σ_s and ϕ_t are state and year fixed effects used to take level difference between states into account and control for unobserved shocks, respectively. Drawing on Bertrand et al. (2004), all SEs are clustered at the state level to allow for correlation of errors over time within states. The coefficients of interest $\beta_{e,k}$ can be interpreted as an average effect of the treatment on the treated periods after initial treatment.

5.6.2 Effects on Police Performance

In this section I move to estimate the impact of raising the felony theft thresholds on police performance by estimating Equation 2. I use the clearance rate, the ratio of arrests to known offenses, as a measure of police performance. Because of the underreported nature of official records, the net effect of raising the felony theft thresholds on clearance is ambiguous, even though a deterioration in police performance is expected. On the one hand, it could cause a fall in the clearance rate. The decline could be driven by fewer arrests and/or more incidence after implementing the higher felony threshold. On the other hand, I may observe an increase in the clearance rate. Decreasing the severity of punishment could discourage victims to report the crime, decrease the recorded incidence, and thereby increase the clearance rate.

The IW estimates will yield a biased estimate of the causal effect of raising the threshold value if downgrading felonies disproportionately occur in states that

[18] Although treatment is defined by the earliest increase in the threshold value during the sample period, in robustness analysis, I also drop states with multiple increments and find similar results in Appendix Figure A5.7.9.

would have experienced a change in the clearance rate even absent the imposition of new thresholds. This would be the case if, for example, states with the clearance burden for felonies strategically wielded some influence on the law change. Although I cannot rule out such a possibility, I can check for differential clearance trends between treatment and control states in the years leading up to a change in the felony theft thresholds.

Figure 5 displays the corresponding estimates. Prior to raising the threshold value, I find little evidence of differential group trends in the theft clearance rate. For $k < 0$, all treatment coefficients are less than 0.01 points in magnitude and never reach statistical significance, supporting the identification assumption.

The effects occur immediately one year after the increase in the felony theft thresholds and are persistent for seven years. Specifically, one year after introducing the higher threshold value, the clearance rates of theft offences decrease significant by 2 percent, and by between 2 and 4 percent in the following four years. Effects then gradually dissipate, reaching insignificance eight years after treatment. The pattern of effects could imply that there is scope of adjustments in police discretion over time following the change in the punitiveness of criminal sanctions. This is consistent with officer learning that police discretion is determined by the experience of officers (West, 2019; DeAngelo and Owens, 2017).

These results complement the finding for crime reporting presented before, suggesting the important impact of the severity of punishments on police behavior. Since the observed negative effects on clearance could be attributable to fewer arrests and/or more incidence, I exploit the impact of adopting higher thresholds for felony theft on apprehension of thieves and incidence of theft, respectively.

5.6.3 Effects on Arrests

According to Lee and McCrary (2017), police officers could exercise discretion in executing an arrest based on the severity of punishment. For instance, a police officer might view a theft with small amounts of cash as forgivable and might overlook the incident upon learning that the it becomes a petty crime after the law change. To shed light on it, I investigate the effects of raising the felony theft thresholds on the logarithm of the number of theft arrests.

As shown in Figure 6, the effects on theft arrests for years prior to the introduction of new thresholds are relatively small and statistically equal to 0, indicating flat pre-trends. Evidently, it shows an effect immediately one year after increasing the threshold value, that is, states that raise the threshold experience a decrease in arrests of thieves than states that do not raise the threshold one year after the new threshold is introduced. The timing of the effects aligns with the effects on the clearance rates.

In contrast, the gap in arrests between them shrinks three years after the treatment. Therefore, the effects of the imposition of higher threshold on the clearance rates are not fully attributable to its effect on police arrests. Also, it is worth noting that the direction of estimates becomes positive and the magnitude of estimates becomes large starting in year 9. They are statistically insignificant, but are suggestive and supportive of possible adjustments in police discretion over time.

Unfortunately, the data on the exact date of state changing the thresholds is not available. Therefore, I am not able to take account of the particular timing during a year when the threshold was raised. As shown in Figure 5 and Figure 6, there are no significant effects of raising felony thresholds on theft clearance and arrests. It is worthy emphasising that lack of exact timing data on threshold changes could potentially explain why I do not detect its effect in the first year of threshold change, as states may raise the thresholds at the end of the enactment year.

5.6.4 Effects on Crime Incidence

Raising the felony theft thresholds could tempt offenders to commit more crimes because offenders would observe less punitive criminal justice sanctions reinforced by a decline in police responsiveness and crime reporting. Particularly, those thieves who previously committed felonies would commit more crimes when they realize that many felonies have been downgraded to misdemeanors following the increase of felony threshold. In this section, I examine the effects of downgrading felonies on the logarithm of the number of recorded theft.

If downgrading felonies results in a fall in official records of theft, the channel behind the effects would be the reduced theft reporting by victims. However, if a rise in theft records is documented, then the effects would be explained by weaker deterrence. Figure 7 presents the corresponding estimates. In the years leading up to the change in the thresholds, police records of theft are statistically identical in level and trend between the states with and without new thresholds. Similarly, moving to the post-period, the effects on theft records appear gradually over time as observed in event studies of the clearance rates and arrests. These positive and significant effects persist for up to seven years: theft records increase significant by 6 percent two years following the threshold change, and by 6 to 12 percent in the following three years; effects then gradually fade.

The absence of a negative impact of higher felony theft thresholds on theft records allows me to interpret these results as evidence that reducing the severity of punishments increases crime incidence via weakening deterrence. Meanwhile, a comparison of the time pattern of the effects on the theft clearance rates with that of the effects on official theft records indicates a one-year delayed response by criminals. This could reflect that it takes time for criminals to learn about the change in police behavior after raising the thresholds, and respond accordingly.

The mean post-treatment estimate can be translated into a elasticity of 0.07: an increase in the value of the felony theft threshold by 135 percent leads to a 7 percent rise in theft offences.[19] Given that downgrading theft felonies could result in greater underreporting of incidents, the estimate likely represents a lower bound of the true elasticity. To place the estimated elasticity in context, I compare it to those in the prior literature on punishment and theft. In terms of the magnitude, my estimated elasticity is very close to those reported in Chalfin and McCrary (2018) and Draca et al. (2011). The authors report elasticities of theft with respect to police staffing

[19]The value of threshold grows by 135 percent on average over the period.

(-0.08) and police deployment (-0.08), respectively.[20] This means that raising the felony theft threshold would have equivalent effects to a 10% decrease in police on theft offences.

A way to corroborate my results is to test whether crimes apart from theft are also influenced by the introduction of higher thresholds too. This could be a key criticism, as this threshold law is specific to theft offences. If I detect a significant impact for other crimes as well, beyond and above theft, then different potential channels could be at work, namely, systematically changes in institutional settings after the threshold change in general. Appendix Figures 5.7.10 and 5.7.11 present the point estimates of the effects of increasing the thresholds on arrests and records of other crimes. The results in suggest that the threshold change has no effect on behavior of police and offenders for other crimes. Hence, there are no generalized effects on law and order following the threshold change that could drive my main results.

5.6.5 Effects on Insurance Searching

Insurance can cover the loss from theft and offer protection. To understand whether the decrease in punishment shifts people's behavior with respect to reconciliation, I examine the effect of increasing felony theft threshold on Google searches for related terms. At least 85% of homeowners in the U.S. have home insurance, and almost all policies cover loss due to theft regardless of the place of incidence (Croll, 2021; National Association of Insurance Commissioners, 2019). I begin by collecting data on Google searches containing the terms "home insurance" by states, which are available for 2004 and later on Google Trends. Figure 8 presents the effect of increasing thresholds on the search volume for "home insurance". The results suggest that the decrease in punishment severity does not trigger a change in protection behavior by enrolling insurance. Then, I examine its effect on the search volume for "phone insurance". It could be a direct measure of insurance search-

[20] The earlier studies also document prison-theft elasticities of -0.28 (Levitt, 1996) and -0.35 (Barbarino and Mastrobuoni, 2014).

ing behavior because mobile phones are one of the most common stolen items. As shown in Figure 9, I do not find effect on searching for phone insurance.

These results can be explained by two contextual factors. First, theft claims are not common. According to the report by Hoel (2021), only 1 percent of filed insurance claims were about theft. Second, concerns on rising premiums prevent the insured from claiming. About 22 percent of survey respondents said that they don't make claims when they have a small loss for fear it will cost them more over time (Blyskal, 2017).

Taken together, my results highlight the salience of the felony theft thresholds in police responses to theft, and criminal decision-making. Consistent with a host of economics of crime research pointing to strong deterrence effects from a threat of punishment, I find evidence that reducing the severity of punishment by raising the felony thresholds would weaken deterrence and generate a boost in crime.

5.7 Conclusion

To the best of my knowledge, this study provides first causal estimates of the effect of felony thresholds on crime reporting. I find that thefts with value just above the felony thresholds are more likely to be reported to police by the victims. This result under a host of alternative specifications is quite consistent, speaking to the strength of the result.

Since there is a long-standing debate over the change in the felony theft thresholds, I also examine the effects of raising the thresholds on police performance and crime incidence. My results indicate that downgrading felonies significantly decreases crime clearance and increases crime incidence. These findings are consistent with the weaker deterrence account that a decrease in an offender's chances of being caught increases crime.

Overall, this study has two important policy implications. First, unlike the literature, which is primarily concerned with relationships between punishment and offenders, this paper considers interactions between punishment and victims, shining lights on underreporting of crime that also limits the deterrent capability of the criminal justice system. Therefore, offenders should not be the only central component of the criminal justice system. Second, the results show that downsizing prisons by reducing the severity of punishment would weaken deterrence and compromise public safety.

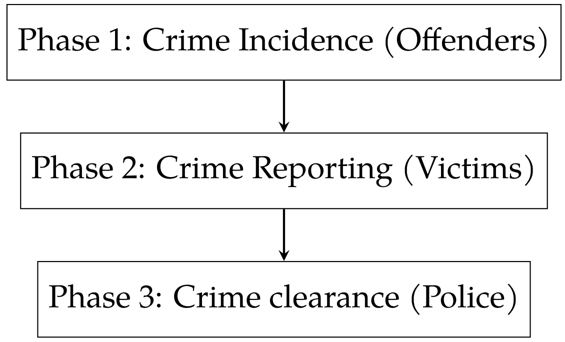


Fig. 1: Crime Situation

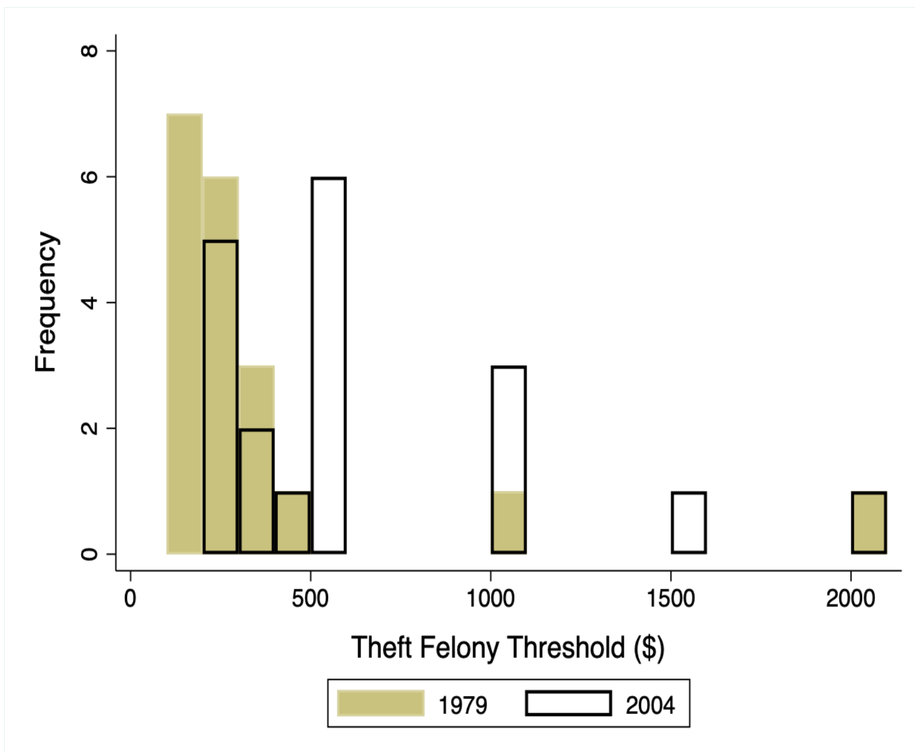


Fig. 2: Distribution of Theft Felony Thresholds

Notes: The figure presents the distribution of the theft felony thresholds for 19 states in the NCVS sample in 1979 and 2004, respectively.

Table 1: Balance Table

	Estimate	S.E.	P-Value	Observations
	(1)	(2)	(3)	(4)
Panel A: Victim				
Female	-0.012	(0.051)	0.83	3,813
White	-0.011	(0.015)	0.46	3,813
Black	-0.006	(0.016)	0.70	3,813
Asian	-0.010	(0.010)	0.34	3,813
Hispanic	-0.024	(0.029)	0.43	3,813
College Education	0.007	(0.078)	0.93	3,823
High Income	0.073	(0.056)	0.21	3,813
Panel B: Incident				
Happen in Commercial Places	-0.009	(0.043)	0.84	3,813
Knew Offenders	0.001	(0.035)	0.98	3,813
Cash on Person	0.027	(0.031)	0.39	3,813
Multiple Offenders	0.000	(0.024)	0.99	3,813

Notes: This table examines whether victim and incident characteristics are balanced across felony threshold. Columns 1 reports the regression discontinuity estimates, following the main estimating equation, of the effect of being above the treatment threshold on these characteristics (with the outcome variable omitted from the set of controls). Column 2 shows the clustered standard errors at the state level in parentheses. Column 3 is the *p*-value for a RD estimate. Column 4 displays the number of observations. An optimal bandwidth of $304 around the threshold has been used.

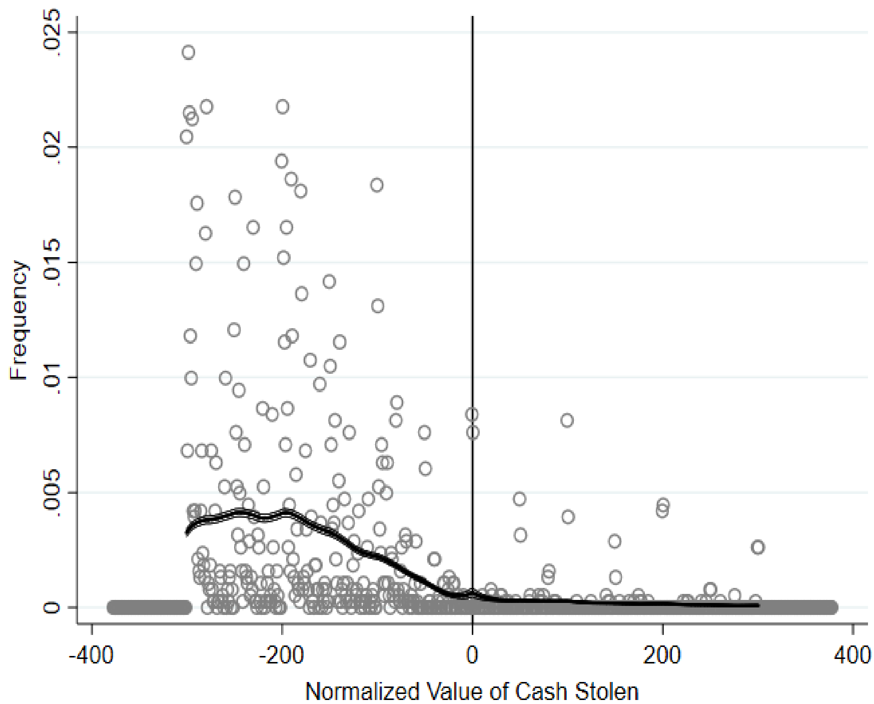


Fig. 3: Density of Cash Theft Value

Notes: The figure shows the density plot, following McCrary (2008). The solid line plots predicted values from a linear regression of frequency on normalized cash theft value, with separate cash theft value trends estimated on either side of the felony threshold. The dashed lines show 95% confidence intervals. The McCrary test statistic is 0.00 (standard error 0.22).

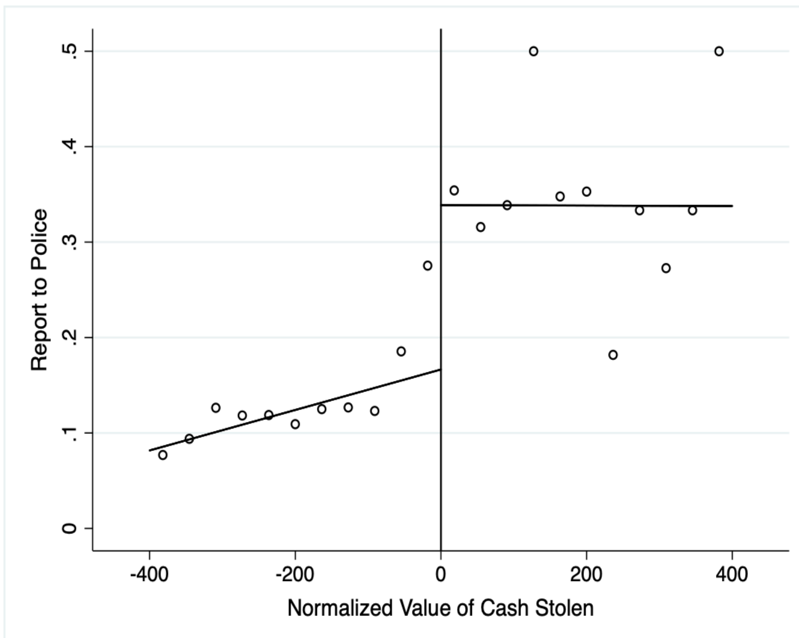


Fig. 4: Effect of Felony Threshold on Crime Reporting

Notes: This figure plots the probability of reporting cash theft to police against the cash theft value. Each point represents the average value of the outcome in the theft value bin. The solid line plots depict first-order polynomial fits on either side of the felony threshold, separately.

Table 2: RD Estimates of the Effect of Felony Threshold on Crime Reporting

Dependent Variable: Was the Crime Reported to Police? (Yes = 1)			
	(1)	(2)	(3)
Felony	0.138***	0.141***	0.141***
	(0.039)	(0.039)	(0.038)
State- and Year- Fixed Effects	YES	YES	YES
Victim Characteristics	NO	YES	YES
Other Controls	NO	NO	YES
Observations	3,813	3,813	3,813

Notes: This table presents RD estimates of the effect of being above the felony theft threshold on crime reporting. Victim characteristics controls include gender, race, Hispanic origin, and indicators for college education, , and high-income household. Other controls include commercial places, acquaintance with the offender, cash on person, and multiple offenders. Standard errors, clustered at the state level, are reported in parenthesis. Significance at the 10%, 5%, and 1% levels is indicated by *, **, and ***, respectively. An optimal bandwidth around the felony theft threshold has been used.

Table 3: RD Estimates of the Effect of Theft Felony Threshold on Crime Reporting for Varying Bandwidths

Dependent Variable: Was the Crime Reported to Police? (Yes = 1)						
	(1)	(2)	(3)	(4)	(5)	(6)
	\$100	\$200	\$300	\$400	\$500	\$600
Felony	0.089*	0.121**	0.202***	0.167***	0.145***	0.160***
	(0.044)	(0.050)	(0.034)	(0.040)	(0.040)	(0.046)
F-statistic	51.6	36.5	28.0	32.82	30.5	25.1
Fixed Effects	YES	YES	YES	YES	YES	YES
Controls	YES	YES	YES	YES	YES	YES
Observations	654	2,073	3,735	4,830	5,212	5,260

Notes: This table presents RD estimates of the effect of being above the felony theft threshold on crime reporting for varying bandwidths. Controls include victim and incident characteristics: gender, race, Hispanic origin, and indicators for college education, high-income household, commercial places, acquaintance with the offender, cash on person, and multiple offenders. Standard errors, clustered at the state level, are reported in parenthesis. Significance at the 10%, 5%, and 1% levels is indicated by *, **, and ***, respectively.

Table 4: Donut RD Estimates of the Effect of Theft Felony Threshold on Crime Reporting

Dependent Variable: Was the Crime Reported to Police? (Yes = 1)				
	(1)	(2)	(3)	(4)
	$1	$10	$50	$100
Felony	0.227***	0.195***	0.222**	0.347**
	(0.055)	(0.061)	(0.088)	(0.150)
Fixed Effects	YES	YES	YES	YES
Controls	YES	YES	YES	YES
Observations	3,752	3,743	3,622	3,159

Notes: This table presents RD estimates of the effect of being above the felony theft threshold on crime reporting when excluding observations within $1, $10, $50, and $100 of the felony threshold, respectively. Controls include victim and incident characteristics: gender, race, Hispanic origin, and indicators for college education, high-income household, commercial places, acquaintance with the offender, cash on person, and multiple offenders. Standard errors, clustered at the state level, are reported in parenthesis. Significance at the 10%, 5%, and 1% levels is indicated by *, **, and ***, respectively.

Table 5: RD Estimates of the Effect of Felony Threshold on Crime Reporting for Rounding Controls

Dependent Variable: Was the Crime Reported to Police? (Yes = 1)			
	(1)	(2)	(3)
Felony	0.200***	0.203***	0.201***
	(0.035)	(0.036)	(0.035)
Multiples of $5	YES	NO	YES
Multiples of $10	NO	YES	YES
State- and Year- Fixed Effects	YES	YES	YES
Other Controls	YES	YES	YES
Observations	3,813	3,813	3,813

Notes: This table presents RD estimates of the effect of being above the felony theft threshold on crime reporting when controlling for multiples of $5 and $10. Other controls include victim and incident characteristics: gender, race, Hispanic origin, and indicators for college education, high-income household, commercial places, acquaintance with the offender, cash on person, and multiple offenders. Standard errors, clustered at the state level, are reported in parenthesis. Significance at the 10%, 5%, and 1% levels is indicated by *, **, and ***, respectively. An optimal bandwidth around the felony theft threshold has been used.

Table 6: RD Estimates of the Effect of Theft Felony Threshold on Reporting of Other Property Crime

Dependent Variable: Was the Crime Reported to Police? (Yes = 1)			
	(1)	(2)	(3)
Felony	-0.097	-0.095	-0.098
	(0.093)	(0.090)	(0.096)
State- and Year- Fixed Effects	YES	YES	YES
Victim Characteristics	NO	YES	YES
Observations	1,397	1,397	1,397

Notes: This table presents RD estimates of the effect of being above the felony theft threshold on crime reporting for other property crime including burglary and robbery. Controls include victim and incident characteristics: gender, race, Hispanic origin, and indicators for college education, high-income household, commercial places, acquaintance with the offender, cash on person, and multiple offenders. Standard errors, clustered at the state level, are reported in parenthesis. Significance at the 10%, 5%, and 1% levels is indicated by *, **, and ***, respectively. An optimal bandwidth $304 has been used.

Table 7: Placebo Tests Using Alternative Thresholds

	(1)	(2)	(3)
Placebo Thresholds (Relative To Felony Thresholds)	Coefficients	Standard Errors	Observations
Panel A ($)			
-299	-0.02	0.03	6,803
+200	0.01	0.15	382
Panel B ($)			
-737	-0.04	0.03	3,340
-197	-0.03	0.02	3,463
+120	-0.06	0.22	194
+549	-0.11	0.13	188

Notes: Notes: The table shows RD estimates of the effect of placebo thresholds on homicides against women for the election subsample. The following theft value ranges (relative to felony threshold values) are used: [-1999, 0), (0, 72411], [-1999, -299), [-299, 0), [0, 200], and (200, 72411]. In each range a placebo RD threshold is created at the median.

Table 8: States Raising the Felony Theft Thresholds (2001-2016)

State	Year of Change
Alaska	2014, 2016
Alabama	2003, 2015
Arkansas	2011
Arizona	2006
California	2010
Colorado	2007, 2013
Connecticut	2009
Delaware	2009
District of Columbia	2010
Georgia	2012
Hawaii	2016
Illinois	2010
Indiana	2013
Kansas	2004, 2016
Kentucky	2009
Louisiana	2009, 2014
Maryland	2009, 2016
Missouri	2002, 2014
Mississippi	2003, 2014
Minnesota	2007
Montana	2009
New Hampshire	2010
New Mexico	2006
Nebraska	2015
Nevada	2011
North Dakota	2013
Oklahoma	2001, 2016
Oregon	2009
Ohio	2011
Rhode Island	2012
South Carolina	2010
South Dakota	2005
Tennessee	2016
Texas	2015
Utah	2010
Vermont	2006
Washington	2009
Wyoming	2004

Notes: This table presents the names of states that raising the felony theft thresholds, and the years of change between 2001 and 2016 in the U.S..

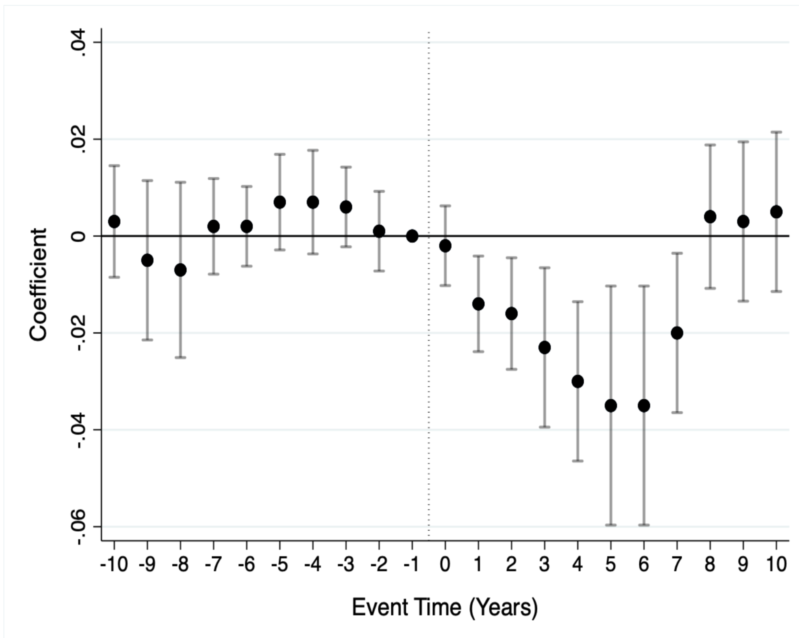


Fig. 5: Event Study Analysis for Theft Clearance Rate

Notes: The figure presents the event-study plots of the effect of the increase of theft felony threshold on theft clearance rate, following Sun and Abraham (2021). Standard errors are clustered at the state level and observations are weighted by state population. Controls include population, percentage of black population, police size, rate of obtaining high school diploma or above by sex, unemployment by sex, income by sex. The dotted vertical line represents the time of treatment. Bands indicate 90 percent confidence intervals.

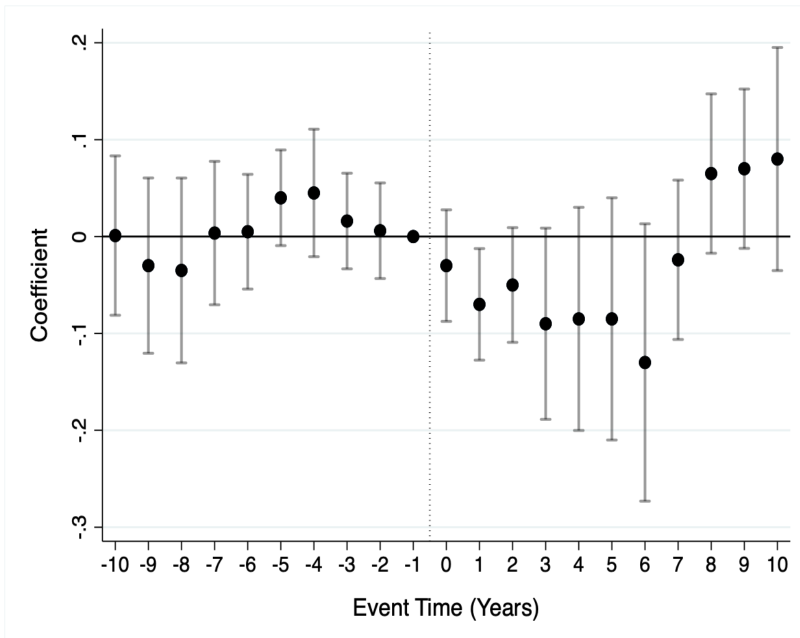


Fig. 6: Event Study Analysis for Police Arresting

Notes: The figure presents the event-study plots of the effect of the increase of theft felony threshold on the logarithm of the counts of thefts cleared by arrest, following Sun and Abraham (2021). Standard errors are clustered at the state level and observations are weighted by state population. Controls include population, percentage of black population, police size, rate of obtaining high school diploma or above by sex, unemployment by sex, income by sex. The dotted vertical line represents the time of treatment. Bands indicate 90 percent confidence intervals.

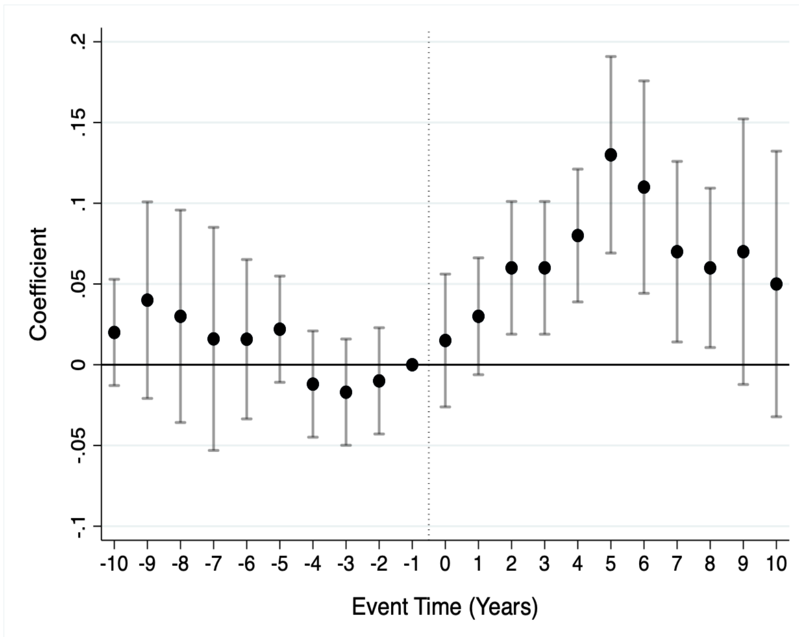


Fig. 7: Event Study Analysis for Theft Incidence

Notes: The figure presents the event-study plots of the effect of the increase of theft felony threshold on the logarithm of the counts of theft incidence, following Sun and Abraham (2021). Standard errors are clustered at the state level and observations are weighted by state population. Controls include population, percentage of black population, police size, rate of obtaining high school diploma or above by sex, unemployment by sex, income by sex. The dotted vertical line represents the time of treatment. Bands indicate 90 percent confidence intervals.

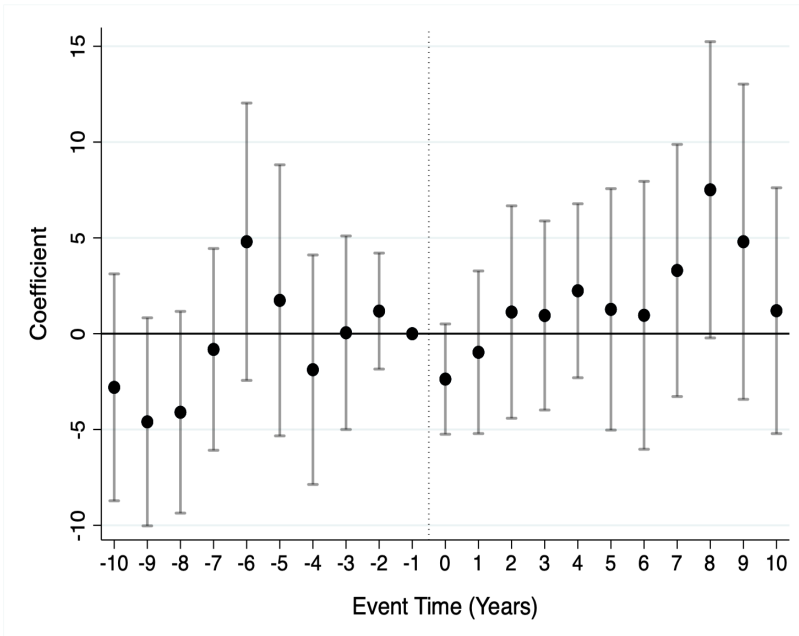


Fig. 8: Event Study Analysis for "Home Insurance" Searches

Notes: The figure presents the event-study plots of the effect of the increase of felony theft threshold on Google search for "home insurance", following Sun and Abraham (2021). Standard errors are clustered at the state level and observations are weighted by state population. Controls include population, percentage of black population, police size, rate of obtaining high school diploma or above by sex, unemployment by sex, income by sex. The dotted vertical line represents the time of treatment. Bands indicate 90 percent confidence intervals.

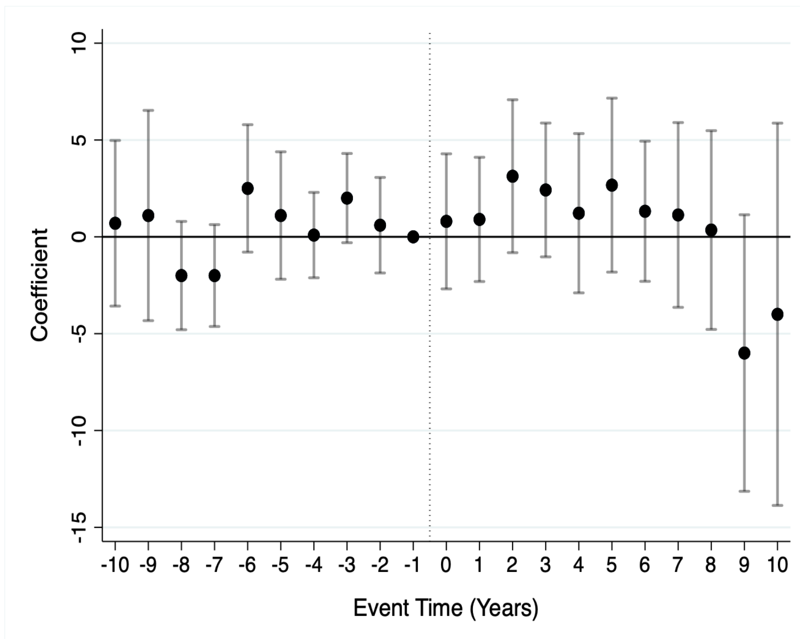


Fig. 9: Event Study Analysis for “Phone Insurance” Searches

Notes: The figure presents the event-study plots of the effect of the increase of felony theft threshold on Google search for “phone insurance”, following Sun and Abraham (2021). Standard errors are clustered at the state level and observations are weighted by state population. Controls include population, percentage of black population, police size, rate of obtaining high school diploma or above by sex, unemployment by sex, income by sex. The dotted vertical line represents the time of treatment. Bands indicate 90 percent confidence intervals.

Table A5.1: RD Estimates of the Effect of Theft Felony Threshold on Crime Reporting for Alternative Polynomials

Dependent Variable: Was the Crime Reported to Police? (Yes = 1)			
	(1)	(2)	(3)
	Linear	Quadratic	Cubic
Felony	0.202***	0.171***	0.129**
	(0.037)	(0.043)	(0.061)
Controls	YES	YES	YES
Optimal Bandwidth	$304	$504	$767
Observations	3,813	5,217	5,565

Notes: This table presents RD estimates of the effect of being above the felony theft threshold on crime reporting for alternative polynomials. Controls include victim and incident characteristics: gender, race, Hispanic origin, and indicators for college education, high-income household, commercial places, acquaintance with the offender, cash on person, and multiple offenders. Standard errors, clustered at the state level, are reported in parenthesis. Significance at the 10%, 5%, and 1% levels is indicated by *, **, and ***, respectively. Optimal bandwidths have been used for alternative RD polynomials, respectively.

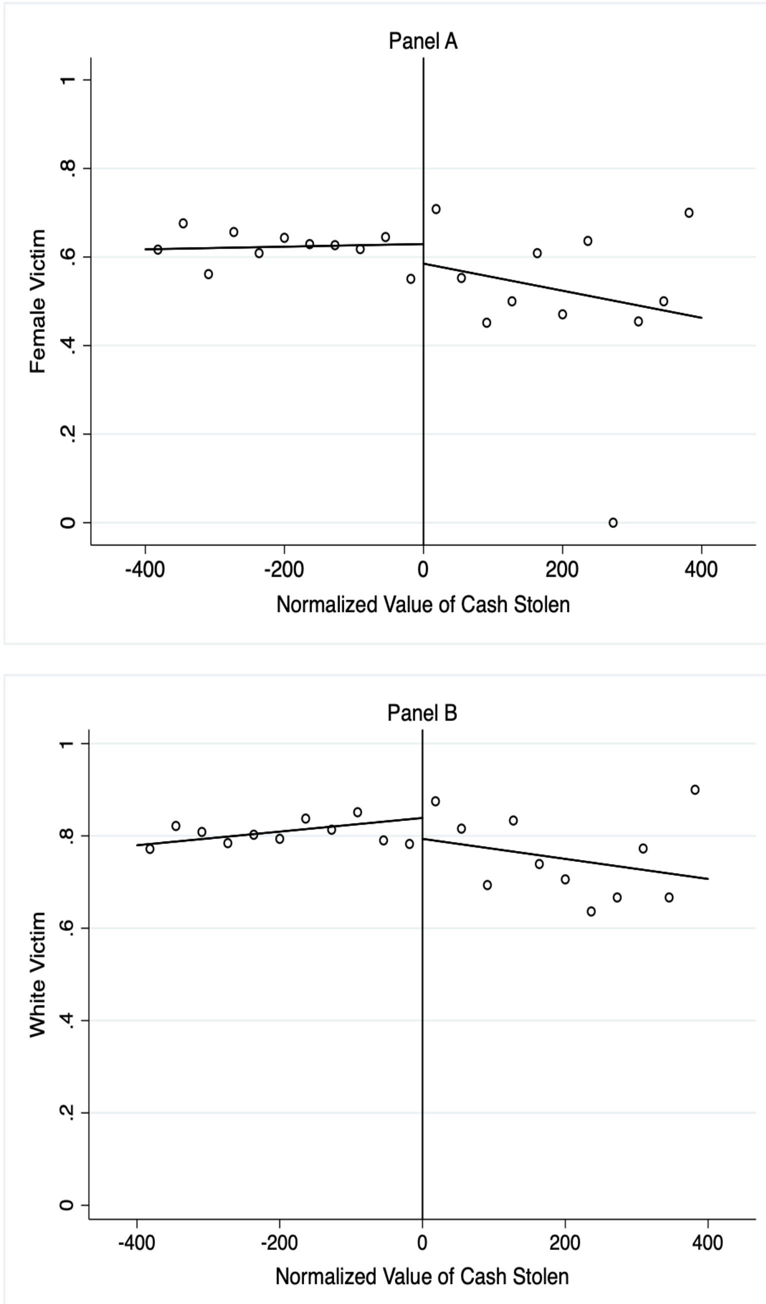


Fig. 10: Appendix Figure 1: Balance of Characteristics

Notes: This figure plots the characteristics against the cash theft value. Points to the right of 0 are above treatment thresholds, while points to the left of 0 are below treatment thresholds. Each point represents the average value of the outcome in the theft value bin. The solid line plots depict first-order polynomial fits on either side of the felony threshold, separately.

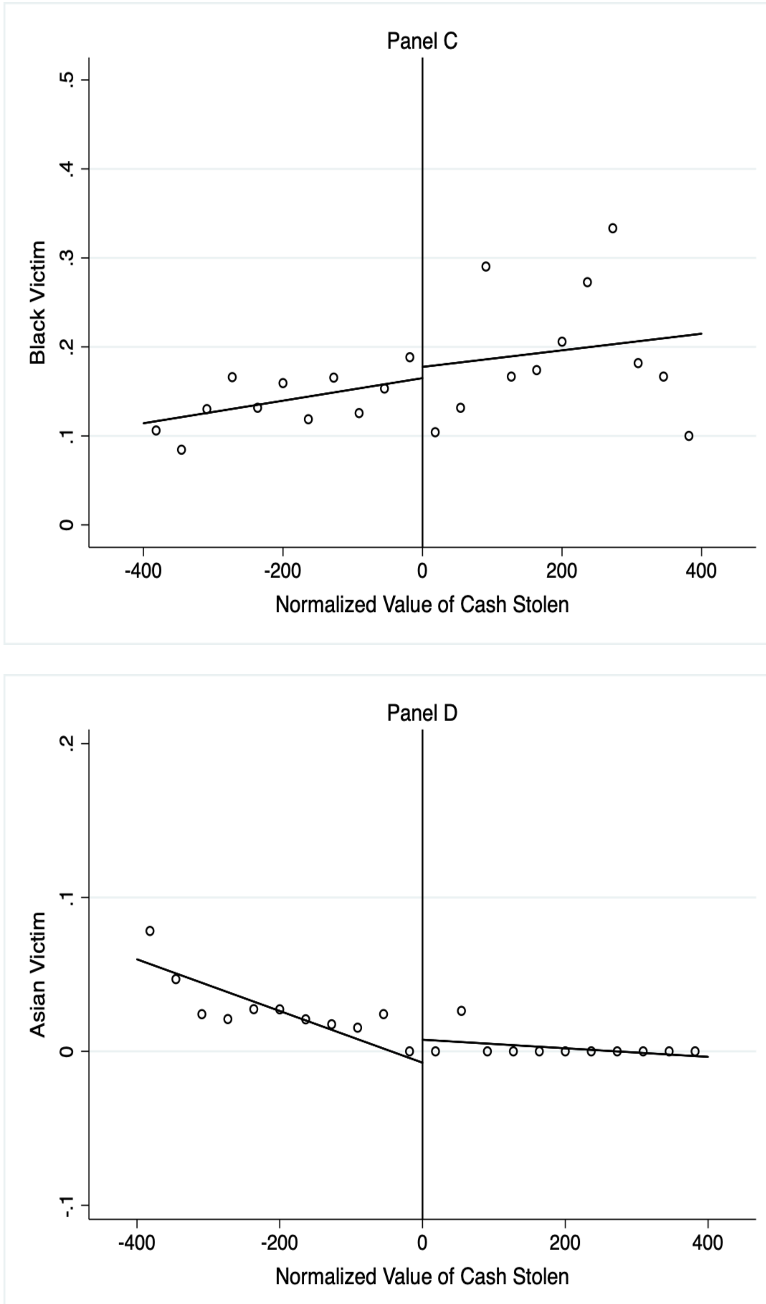
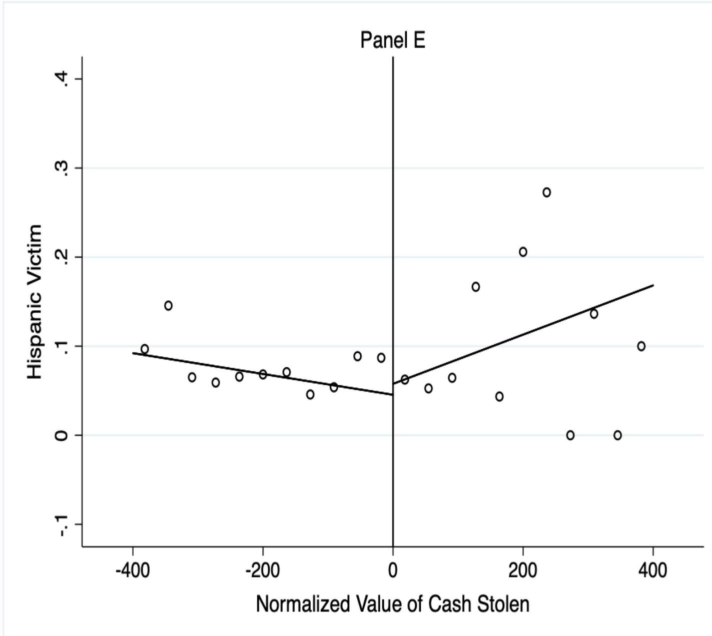


Fig. 11: Appendix Figure 1: Balance of Characteristics (Continued)

Notes: This figure plots the characteristics against the cash theft value. Points to the right of 0 are above treatment thresholds, while points to the left of 0 are below treatment thresholds. Each point represents the average value of the outcome in the theft value bin. The solid line plots depict first-order polynomial fits on either side of the felony threshold, separately.



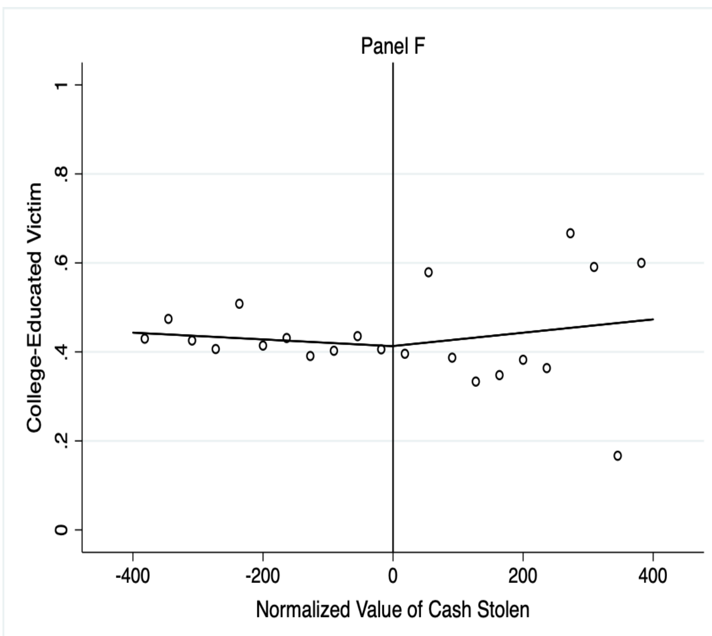


Fig. 12: Appendix Figure 1: Balance of Characteristics (Continued)

Notes: This figure plots the characteristics against the cash theft value. Points to the right of 0 are above treatment thresholds, while points to the left of 0 are below treatment thresholds. Each point represents the average value of the outcome in the theft value bin. The solid line plots depict first-order polynomial fits on either side of the felony threshold, separately.

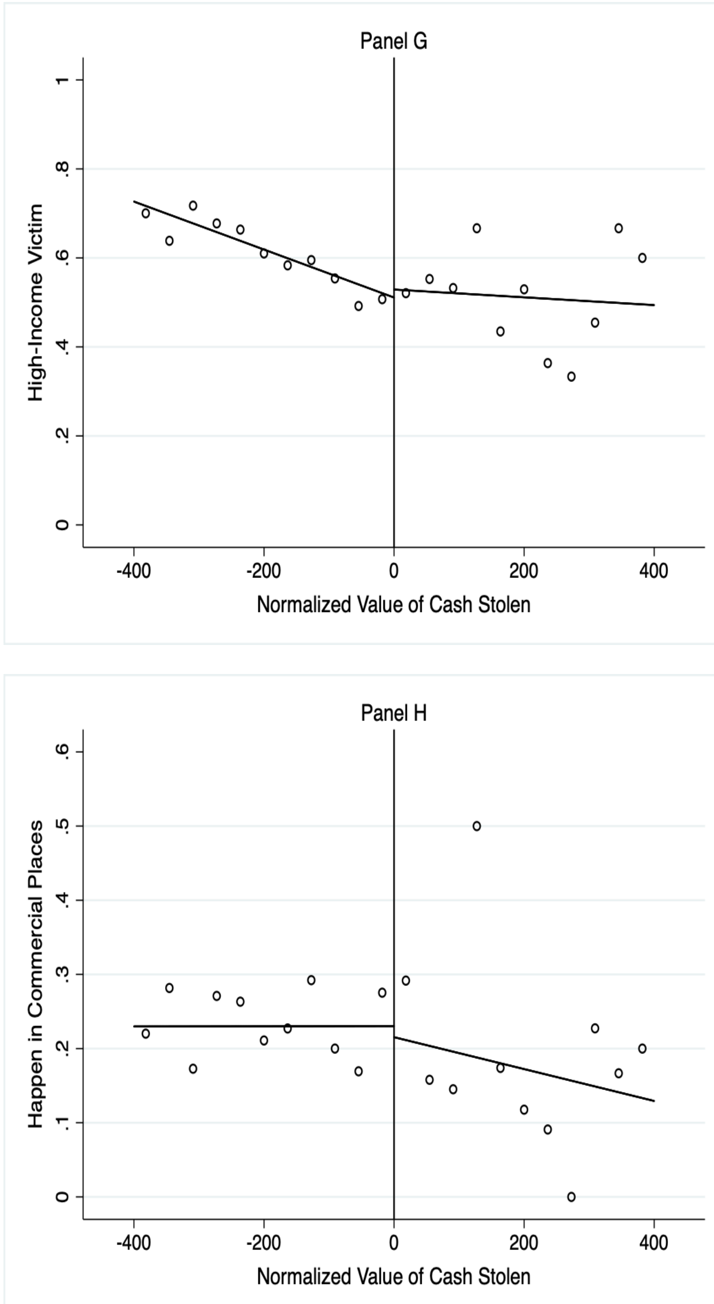
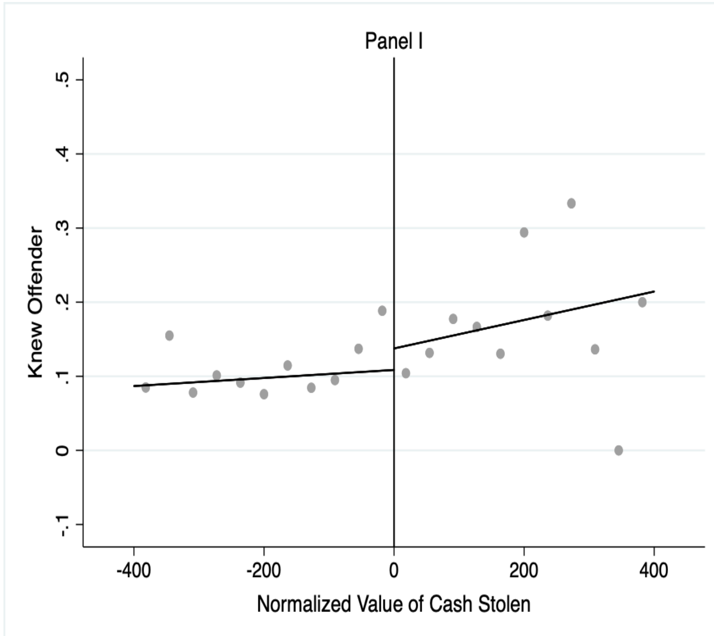


Fig. 13: Appendix Figure 1: Balance of Characteristics (Continued)

Notes: This figure plots the characteristics against the cash theft value. Points to the right of 0 are above treatment thresholds, while points to the left of 0 are below treatment thresholds. Each point represents the average value of the outcome in the theft value bin. The solid line plots depict first-order polynomial fits on either side of the felony threshold, separately.



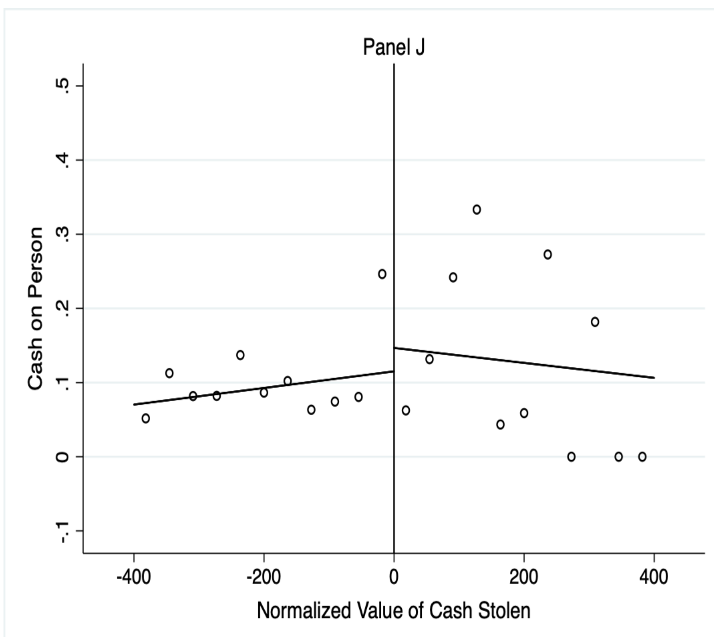


Fig. 14: Appendix Figure 1: Balance of Characteristics (Continued)

Notes: This figure plots the characteristics against the cash theft value. Points to the right of 0 are above treatment thresholds, while points to the left of 0 are below treatment thresholds. Each point represents the average value of the outcome in the theft value bin. The solid line plots depict first-order polynomial fits on either side of the felony threshold, separately.

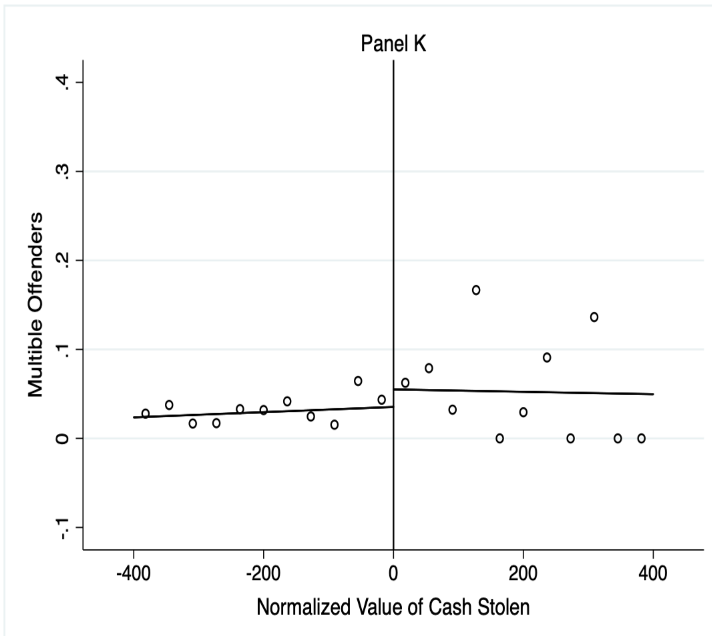


Fig. 15: Appendix Figure 1: Balance of Characteristics (Continued)

Notes: This figure plots the characteristics against the cash theft value. Points to the right of 0 are above treatment thresholds, while points to the left of 0 are below treatment thresholds. Each point represents the average value of the outcome in the theft value bin. The solid line plots depict first-order polynomial fits on either side of the felony threshold, separately.

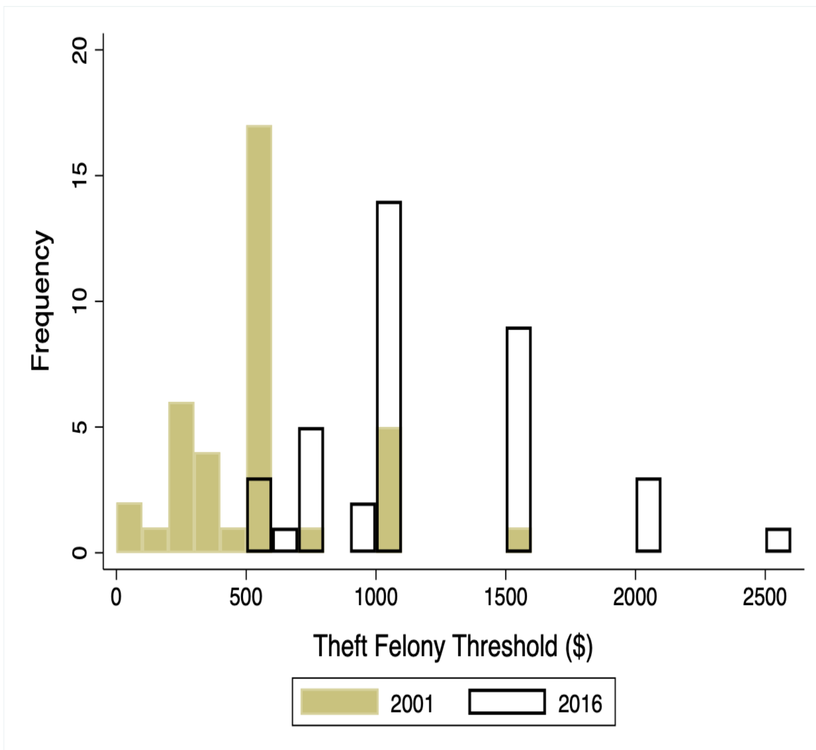


Fig. 16: Distribution of Theft Felony Thresholds

Notes: The figure presents the distribution of the theft felony thresholds in 38 changing states in 2001 and 2016, respectively.

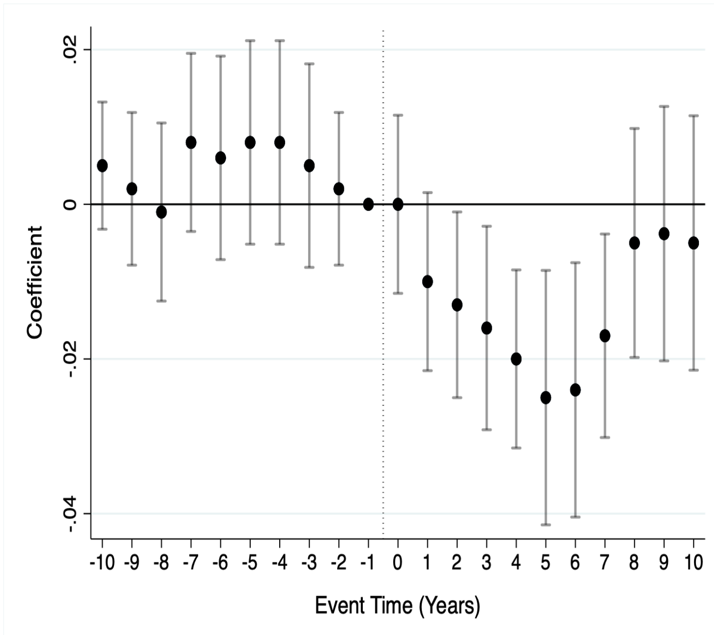


Fig. 17: Event Study Analysis for Theft Clearance Rate (TWFE DD)

Notes: The figure presents the event-study plots of the effect of the increase of theft felony threshold on clearance rate of theft, following standard two-way fixed effect (TWFE) DD model. Standard errors are clustered at the state level and observations are weighted by state population. Controls include population, percentage of black population, police size, rate of obtaining high school diploma or above by sex, unemployment by sex, income by sex. The dotted vertical line represents the time of treatment. Bands indicate 90 percent confidence intervals.

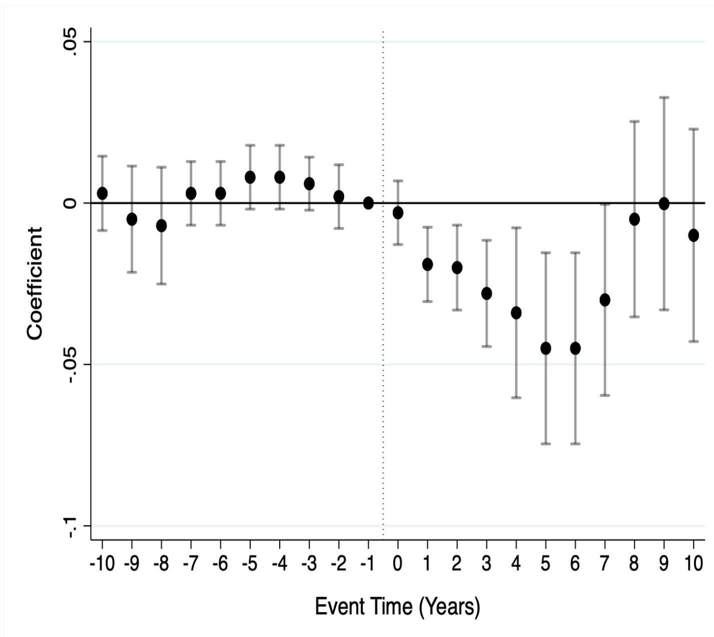


Fig. 18: Event Study Analysis for Theft Clearance Rate

Notes: The figure presents the event-study plots of the effect of the increase of theft felony threshold on clearance rate of theft after excluding states with multiple increments in theft thresholds, following Sun and Abraham (2021). Standard errors are clustered at the state level and observations are weighted by state population. Controls include population, percentage of black population, police size, rate of obtaining high school diploma or above by sex, unemployment by sex, income by sex. The dotted vertical line represents the time of treatment. Bands indicate 90 percent confidence intervals.

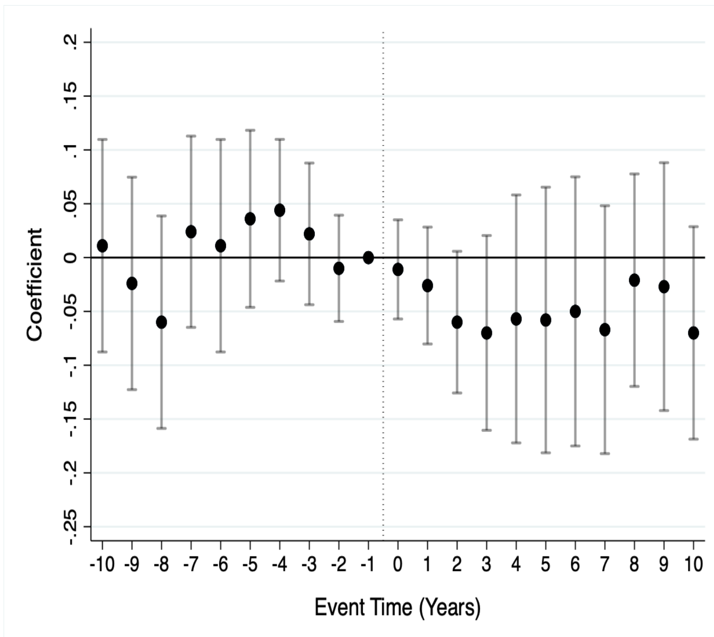


Fig. 19: Event Study Analysis for Arrests of Other Crimes

Notes: The figure presents the event-study plots of the effect of the increase of theft felony threshold on the logarithm of the counts of other crimes cleared by arrest, following Sun and Abraham (2021). Standard errors are clustered at the state level and observations are weighted by state population. Controls include population, percentage of black population, police size, rate of obtaining high school diploma or above by sex, unemployment by sex, income by sex. The dotted vertical line represents the time of treatment. Bands indicate 90 percent confidence intervals.

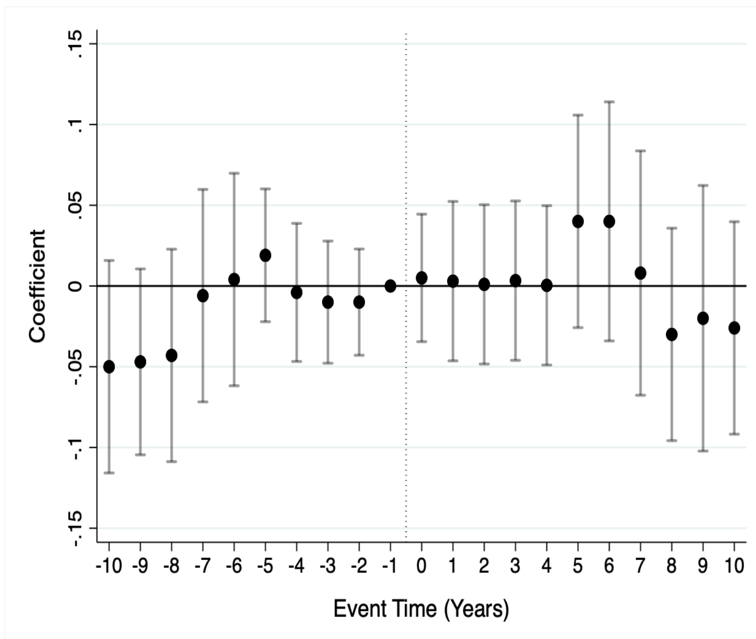


Fig. 20: Event Study Analysis for Other Crime Incidence

Notes: The figure presents the event-study plots of the effect of the increase of theft felony threshold on total crime excluding theft offences, following Sun and Abraham (2021). Standard errors are clustered at the state level and observations are weighted by state population. Controls include population, percentage of black population, police size, rate of obtaining high school diploma or above by sex, unemployment by sex, income by sex. The dotted vertical line represents the time of treatment. Bands indicate 90 percent confidence intervals.

CONCLUSION

This thesis studies two main themes in economics and their applications in the context of equality and justice with a focus on gender and law. Four of the chapters look at gender inequality. The fifth chapter looks at whether the severity of a criminal sanction affects justice and public safety. The five papers centre on different contexts, three in the US, one in the UK, and one in India. This thesis shows that achieving equality and justice remains unfinished business in both developed and developing countries.

In Chapter One, I examine whether electing female mayors can help women in the population. I offer the first comprehensive analysis of the effect of female mayors on violence against women by exploiting the female win-loss threshold.

In Chapter Two, I show whether women's representation in leadership team of law enforcement agencies can contribute to the improved assurance of personal safety for women. I find that women serving in positions of law enforcement leadership significantly lead to a decrease in violencea against women

In Chapter Three, we explore whether exposure to statues can motivate students. Using a RCT, we find that statues have a role modelling effect on students and improve their academic performance.

In Chapter Four, I study whether improved access to health care saves battered women's lives. Using different measures of accessing health care, I find similar evidence that making health care more affordable and accessible helps to save lives of battered women.

In Chapter Five, I investigate whether the punitiveness of sentencing regime influences crime reporting that is of first order importance in production of law enforcement services. My results also indicate that downgrading felonies significantly decreases crime clearance and increases crime incidence

In this thesis I have touched upon both these themes: showing the benefits of empowering women and the cost of reducing the severity of sanction regime. These are important areas which could benefit from a deeper understanding, and are important topics both for researchers, but also policy makers, going forwards.

REFERENCES

Abp (2022). Rajasthan Schools Reopen: Class 1 To 5 Reopen After 2 Years. https://news.abplive.com/news/india/rajasthan-schools-reopen-class-1-to-5-reopen-after-2-years-other-covid-19-guidelines-relaxed-rts-1513350.

Agan, A. Y., Doleac, J. L., and Harvey, A. (2021). Misdemeanor prosecution. Technical report, National Bureau of Economic Research.

Aglasan, S., Guiteras, R., and Palloni, G. (2017). A practitioner's guide to randomization inference. Technical report, Working paper.

Ahern, K. R. and Dittmar, A. K. (2012). The changing of the boards: The impact on firm valuation of mandated female board representation. *The quarterly journal of economics*, 127(1):137–197.

Ahmed, H., Mahmud, M., Said, F., and Tirmazee, Z. S. (2022). Encouraging female graduates to enter the labor force: Evidence from a role model intervention in pakistan.

Aish, N., Asare, P., and Miskioğlu, E. E. (2018). People like me: Providing relatable and realistic role models for underrepresented minorities in stem to increase their motivation and likelihood of success. In *2018 IEEE integrated STEM education conference (ISEC)*, pages 83–89. IEEE.

Aizer, A. (2010). The Gender Wage Gap and Domestic Violence. *American Economic Review*, 100(4):1847–59.

Aizer, A. (2011). Poverty, Violence, and Health the Impact of Domestic Violence During Pregnancy on Newborn Health. *Journal of Human resources*, 46(3):518–538.

Aizer, A. and Dal Bo, P. (2009). Love, Hate and Murder: Commitment Devices in Violent Relationships. *Journal of Public Economics*, 93(3-4):412–428.

Akerlof, G. A. and Kranton, R. E. (2000). Economics and identity. *The quarterly journal of economics*, 115(3):715–753.

Alesina, A., Brioschi, B., and La Ferrara, E. (2016). Violence against women: A cross-cultural analysis for africa. Technical report, National Bureau of Economic Research.

Alicke, M. D. and Yurak, T. J. (1995). Perpetrator personality and judgments of acquaintance rape. *Journal of Applied Social Psychology*, 25(21):1900–1921.

Allen, W. D. (2007). The reporting and underreporting of rape. *Southern Economic Journal*, pages 623–641.

Amaral, S., Bhalotra, S., and Prakash, N. (2021). Gender, crime and punishment: Evidence from women police stations in india.

Anastasopoulos, L. (2016). Estimating the gender penalty in house of representative elections using a regression discontinuity design. *Electoral Studies*, 43:150–157.

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

Anderson, E. and Clibbens, P. (2018). 'smugglers of truth': The indian diaspora, hindu nationalism, and the emergency (1975–77). *Modern Asian Studies*, 52(5):1729–1773.

Anderson, K. B., Cooper, H., and Okamura, L. (1997). Individual Differences and Attitudes toward Rape: A Meta-analytic Review. *Personality and Social Psychology Bulletin*, 23(3):295–315.

Angelucci, M. (2007). Love on the rocks: Alcohol abuse and domestic violence in rural mexico.

Antonovics, K. and Knight, B. G. (2009). A new look at racial profiling: Evidence from the boston police department. *The Review of Economics and Statistics*, 91(1):163–177.

Anwar, S., Bayer, P., and Hjalmarsson, R. (2019). A Jury of Her Peers: The Impact of the First Female Jurors on Criminal Convictions. *The Economic Journal*, 129(618):603–650.

Aral, S. and Walker, D. (2012). Identifying influential and susceptible members of social networks. *Science*, 337(6092):337–341.

Ariely, D., Gneezy, U., Loewenstein, G., and Mazar, N. (2009). Large stakes and big mistakes. *The Review of Economic Studies*, 76(2):451–469.

Asgari, S., Dasgupta, N., and Cote, N. G. (2010). When does contact with successful ingroup members change self-stereotypes? a longitudinal study comparing the effect of quantity vs. quality of contact with successful individuals. *Social Psychology*, 41(3):203.

Ashraf, Q. and Galor, O. (2013). The'out of africa'hypothesis, human genetic diversity, and comparative economic development. *American Economic Review*, 103(1):1–46.

Athey, S., Avery, C., and Zemsky, P. (2000). Mentoring and Diversity. *American Economic Review*, 90(4):765–786.

Athey, S. and Imbens, G. W. (2021). Design-based analysis in difference-in-differences settings with staggered adoption. *Journal of Econometrics*.

Ayalew, S., Manian, S., and Sheth, K. (2018). Discrimination from below: Experimental evidence on female leadership in ethiopia. Technical report, Technical report, mimeo.

Bachman, R. (1998). The factors related to rape reporting behavior and arrest: New evidence from the national crime victimization survey. *Criminal justice and behavior*, 25(1):8–29.

Baker, W., Hutchinson, J., Moore, D., and Nedungadi, P. (1986). Brand familiarity and advertising: effects on the evoked set and brand preference. *ACR North American Advances*.

Balkin, J. (1988). Why Policemen Don't Like Policewomen. *Journal of Police Science & Administration*, 16(1):29–38.

Banerjee, A., Chattopadhyay, R., Duflo, E., Keniston, D., and Singh, N. (2012). Improving police performance in rajasthan, india: Experimental evidence on incentives, managerial autonomy and training. Technical report, National Bureau of Economic Research.

Barati, M. (2019). Punishment severity and crime: The case of arkansas. *Review of Law & Economics*, 15(1).

Barbarino, A. and Mastrobuoni, G. (2014). The incapacitation effect of incarceration: Evidence from several italian collective pardons. *American Economic Journal: Economic Policy*, 6(1):1–37.

Bardhan, P. (2002). Decentralization of governance and development. *Journal of Economic perspectives*, 16(4):185–205.

Bardhan, P. K., Mookherjee, D., and Torrado, M. P. (2010). Impact of Political Reservations in West Bengal Local Governments on Anti-poverty Targeting. *Journal of Globalization and Development*, 1(1).

Barreca, A. I., Guldi, M., Lindo, J. M., and Waddell, G. R. (2011). Saving babies? revisiting the effect of very low birth weight classification. *The Quarterly Journal of Economics*, 126(4):2117–2123.

Bartos, B. J. and Kubrin, C. E. (2018). Can we downsize our prisons and jails without compromising public safety? findings from california's prop 47. *Criminology & Public Policy*, 17(3):693–715.

Beaman, L., Chattopadhyay, R., Duflo, E., Pande, R., and Topalova, P. (2009). Powerful Women: Does Exposure Reduce Bias? *The Quarterly Journal of Economics*, 124(4):1497–1540.

Beaman, L., Duflo, E., Pande, R., and Topalova, P. (2012). Female leadership raises aspirations and educational attainment for girls: A policy experiment in india. *science*, 335(6068):582–586.

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

Belknap, J. (2010). Rape: Too hard to report and too easy to discredit victims. *Violence Against Women*, 16(12):1335–1344.

Bell, B., Jaitman, L., and Machin, S. (2014). Crime deterrence: Evidence from the london 2011 riots. *The Economic Journal*, 124(576):480–506.

Bell, S. (2008). Introduction: Role models in the roman world. *Memoirs of the American Academy in Rome. Supplementary Volumes*, 7:1–39.

Benhamou, F. (2020). Heritage. In *Handbook of Cultural Economics, Third Edition*. Edward Elgar Publishing.

Benjamini, Y., Krieger, A. M., and Yekutieli, D. (2006). Adaptive linear step-up procedures that control the false discovery rate. *Biometrika*, 93(3):491–507.

Bennett, T. (2002). Reproductive health care in the rural united states. *JAMA*, 287(1):112–112.

Benson, B. L., Rasmussen, D. W., and Kim, I. (1998). Deterrence and Public Policy: Trade-offs in the Allocation of Police Resources. *International Review of Law and Economics*, 18(1):77–100.

Berchick, E. R., Hood, E., and Barnett, J. C. (2019). *Health insurance coverage in the United States: 2018*. Washington, DC: US Department of Commerce.

Bernard, T., Dercon, S., Orkin, K., Taffesse, A., et al. (2014). *The future in mind: Aspirations and forward-looking behaviour in rural Ethiopia*. Centre for Economic Policy Research London.

Bertrand, M., Black, S. E., Jensen, S., and Lleras-Muney, A. (2019). Breaking the glass ceiling? the effect of board quotas on female labour market outcomes in norway. *The Review of Economic Studies*, 86(1):191–239.

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.

Besley, T., Pande, R., and Rao, V. (2012). Just Rewards? Local Politics and Public Resource Allocation in South India. *The World Bank Economic Review*, 26(2):191–216.

Bettinger, E., Ludvigsen, S., Rege, M., Solli, I. F., and Yeager, D. (2018). Increasing perseverance in math: Evidence from a field experiment in norway. *Journal of Economic Behavior & Organization*, 146:1–15.

Bhalotra, S., Britto, D., Pinotti, P., and Sampaio, B. (2020). Job displacement, unemployment benefits and domestic violence. *Unpublished Manuscript*, 4.

Bhan, P. C. (2020). Do role models increase student hope and effort? evidence from india.

Bharadwaj, P., Løken, K. V., and Neilson, C. (2013). Early life health interventions and academic achievement. *American Economic Review*, 103(5):1862–91.

Bhuller, M., Dahl, G. B., Løken, K. V., and Mogstad, M. (2020). Incarceration, recidivism, and employment. *Journal of Political Economy*, 128(4):1269–1324.

Bjorvatn, K., Cappelen, A. W., Sekei, L. H., Sørensen, E. Ø., and Tungodden, B. (2020). Teaching through television: Experimental evidence on entrepreneurship education in tanzania. *Management Science*, 66(6):2308–2325.

Blackman, A. M. (1923). The sun-cult in ancient egypt. *Nature*, 111(2789):499–502.

Blais, A., Blake, D., and Dion, S. (1993). Do parties make a difference? parties and the size of government in liberal democracies. *American Journal of Political Science*, pages 40–62.

Blavin, F., Karpman, M., Kenney, G. M., and Sommers, B. D. (2018). Medicaid versus marketplace coverage for near-poor adults: effects on out-of-pocket spending and coverage. *Health affairs*, 37(2):299–307.

Blesse, S. and Diegmann, A. (2022). The place-based effects of police stations on crime: Evidence from station closures. *Journal of Public Economics*, 207:104605.

Blyskal, J. (2017). What You Don't Know About Homeowners Insurance Could Cost You. Retrieved from https://www.consumerreports.org/homeowners-insurance/what-you-dont-know-about-homeowners-insurance-could-cost-you/ Accessed 12 Jan 2022.

Boateng, F. D. (2018). Crime reporting behavior: do attitudes toward the police matter? *Journal of interpersonal violence*, 33(18):2891–2916.

Bobonis, G. J., González-Brenes, M., and Castro, R. (2013). Public transfers and domestic violence: The roles of private information and spousal control. *American Economic Journal: Economic Policy*, 5(1):179–205.

Borgschulte, M. and Vogler, J. (2020). Did the aca medicaid expansion save lives? *Journal of Health Economics*, 72:102333.

Bower, G. H. (1970). Analysis of a mnemonic device: Modern psychology uncovers the powerful components of an ancient system for improving memory. *American Scientist*, 58(5):496–510.

Brocas, I. and Carrillo, J. D. (2020). Introduction to special issue “understanding cognition and decision making by children.” studying decision-making in children: Challenges and opportunities. *Journal of Economic Behavior & Organization*, 179:777–783.

Brollo, F. and Troiano, U. (2016). What happens when a woman wins an election? evidence from close races in brazil. *Journal of Development Economics*, 122:28–45.

Brown-Schmidt, S. and Duff, M. C. (2016). Memory and common ground processes in language use.

Buettgens, M., Garrett, B., and Holahan, J. (2010). The Effects of Large Premium Increases on Individuals, Families, and Small Businessesa. Retrieved from https://www.urban.org/sites/default/files/publication/28576/412079-The-Effects-of-Large-Premium-Increases-on-Individuals-Families-and-Small-Businesses.PDF.

Bulman, G. (2019). Law Enforcement Leaders and the Racial Composition of Arrests. *Economic Inquiry*, 57(4):1842–1858.

Buonanno, P. and Raphael, S. (2013). Incarceration and incapacitation: Evidence from the 2006 italian collective pardon. *American Economic Review*, 103(6):2437–65.

Bureau of Justice Statistics (1987). *Felony Laws of the 50 States and the District of Columbia, 1986*. US Department of Justice, Bureau of Justice Statistics.

Bureau of Justice Statistics (2012). Victimizations Not Reported. Retrieved from https://bjs.ojp.gov/content/pub/pdf/vnrp0610.pdf Accessed 12 August 2020.

Business Standard (2021). Rajasthan govt releases fresh guidelines for Lockdown. Retrieved from https://www.business-standard.com/article/current-affairs/rajasthan-imposes-lockdown-from-may-10-to-24-amid-surge-in-covid-19-cases-121050700063$_1.html$.

Callaway, B. and Sant'Anna, P. H. (2021). Difference-in-differences with multiple time periods. *Journal of Econometrics*, 225(2):200–230.

Calonico, S., Cattaneo, M. D., and Titiunik, R. (2014). Robust nonparametric confidence intervals for regression-discontinuity designs. *Econometrica*, 82(6):2295–2326.

Caprioli, M. and Boyer, M. A. (2001). Gender, violence, and international crisis. *Journal of Conflict Resolution*, 45(4):503–518.

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 Giuliano, L. (2014). Does gifted education work? for which students? Technical report, National Bureau of economic research.

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

Catalano, S. M. (2010). *Victimization during household burglary*. US Department of Justice, Office of Justice Programs, Bureau of Justice ....

Cebula, C. and Salsberg, A. (2019). Uncsw.

Census Bureau (2021). 2021 state government tax tables.

Center for American Women and Politics (2017). Women in the U.S. Congress 2017. Retrieved from https://cawp.rutgers.edu/women-us-congress-2017 Accessed 22 May 2020.

Centers for Disease Control (2003). Costs of intimate partner violence against women in the united states. Retrieved from https://www.cdc.gov/violenceprevention/pdf/ipvbook-a.pdf Accessed 12 August 2020.

Centers for Disease Control (2012). Financial burden of medical care: early release of estimates from the national health interview survey, january-june 2011. *Atlanta (GA): Centers of Disease Control and Prevention*.

Centers for Disease Control (2015). Patterns of Health Insurance Coverage Around the Time of Pregnancy Among Women. Retrieved from hhttps://www.cdc.gov/mmwr/preview/mmwrhtml/ss6404a1.htm Accessed 12 August 2020.

Centers for Disease Control (2018a). Personal health care expenditures. Retrieved from https://www.cdc.gov/nchs/data/hus/2019/046-508.pdf Accessed 21 August 2020.

Centers for Disease Control (2018b). The national intimate partner and sexual violence survey: 2015 data brief–updated release. Retrieved from https://www.cdc.gov/violenceprevention/pdf/2015data-brief508.pdf Accessed 11 May 2020.

Centers for Disease Control (2019). Demographic Variation in Health Insurance Coverage: United States. Retrieved from https://www.cdc.gov/nchs/data/nhsr/nhsr159-508.pdf Accessed 2 August 2020.

Centers for Medicare and Medicaid Services (2014). Overview of Medicaid. Retrieved from https://www.medicaid.gov/state-resource-center/mac-learning-collaboratives/learning-collaborative-state-toolbox/downloads/cost-sharing-premium-requirements.pdf Accessed 2 January 2020.

Centers for Medicare and Medicaid Services (2015). 2015 Actuarial Report. Retrieved from https://www.medicaid.gov/medicaid/downloads/medicaid-actuarial-report-2015.pdf Accessed 12 March 2020.

Chalfin, A., Danagoulian, S., and Deza, M. (2021). Covid-19 has strengthened the relationship between alcohol consumption and domestic violence. Technical report, National Bureau of Economic Research.

Chalfin, A. and McCrary, J. (2017). Criminal deterrence: A review of the literature. *Journal of Economic Literature*, 55(1):5–48.

Chalfin, A. and McCrary, J. (2018). Are US Cities Underpoliced? Theory and Evidence. *Review of Economics and Statistics*, 100(1):167–186.

Chattopadhyay, R. and Duflo, E. (2004). Women as Policy Makers: Evidence from a Randomized Policy Experiment in India. *Econometrica*, 72(5):1409–1443.

Chin, A., Couper, M. P., and Beckett, D. (2021). Attrition in a longitudinal online randomized controlled trial with prospective homeowners. *Field Methods*, 33(1):3–24.

Chong, A. and Ferrara, E. L. (2009). Television and divorce: Evidence from brazilian novelas. *Journal of the European Economic Association*, 7(2-3):458–468.

Chung, K.-S. (2000). Role models and arguments for affirmative action. *American Economic Review*, 90(3):640–648.

Ciacci, R. and Sviatschi, M. M. (2021). The effect of adult entertainment establishments on sex crime: evidence from new york city. *The Economic Journal*.

Clarke, R. V. and Harris, P. M. (1992). Auto theft and its prevention. *Crime and Justice*, 16:1–54.

Clots-Figueras, I. (2011). Women in Politics: Evidence from the Indian States. *Journal of Public Economics*, 95(7-8):664–690.

Clots-Figueras, I. (2012). Are female leaders good for education? evidence from india. *American Economic Journal: Applied Economics*, 4(1):212–44.

CNN (2021a). COVID-19 and children in India. https://www.unicef.org/india/coronavirus/covid-19/covid-19-and-children.

CNN (2021b). The Delta variant leaves children vulnerable to Covid-19. https://edition.cnn.com/2021/06/30/health/delta-variant-children-leana-wen-wellness/index.html.

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(1):25–49.

Cohen, W. (1989). Symbols of power: statues in nineteenth-century provincial france. *Comparative Studies in Society and History*, 31(3):491–513.

College of Policing (2018). College of Policing: Guidance for Appointing Chief Officers. Retrieved from https://www.college.police.uk/What-we-do/Support/Health-safety/Documents/6.3716$_{G}uidance_{f}or_{a}ppointment_{v}18.pdf Accessed 11 May 2019.$

Combs Thorsen, V., Sundby, J., and Malata, A. (2012). Piecing together the maternal death puzzle through narratives: the three delays model revisited. *PloS one*, 7(12):e52090.

Copes, H. (1999). Routine activities and motor vehicle theft: A crime specific approach. *Journal of Crime and Justice*, 22(2):125–146.

Cowper-Coles, M. (2020). Women political leaders: the impact of gender on democracy. *The Global Institute for Women's Leadership*.

Crawford, R. and Disney, R. (2018). Wage Regulation and the Quality of Police Applicants. *Economica*, 85(340):701–734.

Croll, M. (2021). Home Insurance Facts and Statistics. Retrieved from https://www.valuepenguin.com/home-insurance-statisticspurchased Accessed 11 Jan 2022.

Crown Prosecution Service (2018). The Code for Crown Prosecutors. Retrieved from https://www.cps.gov.uk/publication/code-crown-prosecutors Accessed 11 June 2019.

Culver, W. L. (1969). Legislature of the state of california 1969 regular session. Technical report, Senate of the State of California.

Current Population Census (2019). Ipums cps. Retrieved from https://www.ipums.org Accessed 21 October 2021.

Dalton, P. S., Ghosal, S., and Mani, A. (2016). Poverty and aspirations failure. *The Economic Journal*, 126(590):165–188.

Dalton, P. S., Rüschenpöhler, J., Uras, B., and Zia, B. (2021). Curating local knowledge: Experimental evidence from small retailers in indonesia. *Journal of the European Economic Association*, 19(5):2622–2657.

Daniel, H., Erickson, S. M., and Bornstein, S. S. (2018). Women's health policy in the united states: an american college of physicians position paper. *Annals of internal medicine*, 168(12):874–875.

Davies, G. (2008). Portrait statues as models for gender roles in roman society. *Memoirs of the American Academy in Rome. Supplementary Volumes*, 7:207–220.

DeAngelo, G. and Owens, E. G. (2017). Learning the ropes: General experience, task-specific experience, and the output of police officers. *Journal of Economic Behavior & Organization*, 142:368–377.

Deza, M., Maclean, J. C., and Solomon, K. (2022). Local access to mental healthcare and crime. *Journal of Urban Economics*, 129:103410.

Dhar, D., Jain, T., and Jayachandran, S. (2022). Reshaping adolescents' gender attitudes: Evidence from a school-based experiment in india. *American Economic Review*, 112(3):899–927.

Di Tella, R. and Schargrodsky, E. (2004). Do police reduce crime? estimates using the allocation of police forces after a terrorist attack. *American Economic Review*, 94(1):115–133.

Dobbie, W., Goldin, J., and Yang, C. S. (2018). The effects of pretrial detention on conviction, future crime, and employment: Evidence from randomly assigned judges. *American Economic Review*, 108(2):201–40.

Dolan, J. (2000). The senior executive service: Gender, attitudes, and representative bureaucracy. *Journal of Public Administration Research and Theory*, 10(3):513–530.

Doleac, J. L. (2017). The effects of dna databases on crime. *American Economic Journal: Applied Economics*, 9(1):165–201.

Doleac, J. L. (2020). Encouraging desistance from crime. *Available at SSRN*.

Draca, M., Koutmeridis, T., and Machin, S. (2019). The changing returns to crime: do criminals respond to prices? *The Review of Economic Studies*, 86(3):1228–1257.

Draca, M. and Machin, S. (2015). Crime and economic incentives. *economics*, 7(1):389–408.

Draca, M., Machin, S., and Witt, R. (2011). Panic on the Streets of London: Police, Crime, and the July 2005 Terror Attacks. *American Economic Review*, 101(5):2157–81.

Du Mont, J. and Myhr, T. L. (2000). So few convictions: The role of client-related characteristics in the legal processing of sexual assaults. *Violence Against Women*, 6(10):1109–1136.

Dube, O. and Harish, S. (2020). Queens. *Journal of Political Economy*, 128(7):2579–2652.

Duflo, E. (2012). Women Empowerment and Economic Development. *Journal of Economic literature*, 50(4):1051–79.

Dunning, T. and Nilekani, J. (2013). Ethnic Quotas and Political Mobilization: Caste, Parties, and Distribution in Indian Village Councils. *American political Science review*, 107(1):35–56.

Edlund, L. and Pande, R. (2002). Why have women become left-wing? the political gender gap and the decline in marriage. *The Quarterly Journal of Economics*, 117(3):917–961.

Egerö, B. and Hammarskjold, M. (1994). Understanding reproductive change: Kenya, tamil nadu, punjab, costa rica.

Eibner, C. and Nowak, S. (2016). *Evaluating the CARE Act: Implications of a Proposal to Repeal and Replace the Affordable Care Act*. Commonwealth Fund.

Erten, B. and Keskin, P. (2018). For better or for worse?: Education and the prevalence of domestic violence in turkey. *American Economic Journal: Applied Economics*, 10(1):64–105.

Espey, E., Dennis, A., and Landy, U. (2019). The importance of access to comprehensive reproductive health care, including abortion: a statement from women's health professional organizations. *American journal of obstetrics and gynecology*, 220(1):67–70.

Eswaran, M. and Malhotra, N. (2011). Domestic violence and women's autonomy in developing countries: theory and evidence. *Canadian Journal of Economics*, 44(4):1222–1263.

Evans, W. N. and Owens, E. G. (2007). COPS and Crime. *Journal of Public Economics*, 91(1-2):181–201.

Federal Bureau of Investigation (2019). The Nation's Two Crime Measures. Retrieved from https://ucr.fbi.gov/crime-in-the-u.s/2019/crime-in-the-u.s.-2019/topic-pages/nations-two-crime-measures Accessed 11 Sep 2021.

Federal Bureau of Investigation (2020). UCR. Retrieved from https://www.fbi.gov/services/cjis/ucr/ Accessed 12 September 2020.

Ferreira, F. and Gyourko, J. (2009). Do political parties matter? evidence from us cities. *The Quarterly journal of economics*, 124(1):399–422.

Ferreira, F. and Gyourko, J. (2014). Does gender matter for political leadership? the case of us mayors. *Journal of Public Economics*, 112:24–39.

Field, E., Pande, R., Rigol, N., Schaner, S., and Moore, C. T. (2016). On her account: Can strengthening women's financial control boost female labor supply? *Harvard University Working Paper*.

Flatley, J. (2018). Sexual offences in England and Wales: Year ending March 2017. Retrieved from https://www.ons.gov.uk/peoplepopulationandcommunity/crimeandjustice/articles/sexualof Accessed 12 June 2019.

Forbesindia (2020). Why youth prefer government jobs over private jobs for building up career. https://www.forbesindia.com/article/brand-connect/heres-why-youth-prefer-government-jobs-over-private-jobs-for-building-up-career/65159/1.

Friedemann-Sánchez, G. and Lovatón, R. (2012). Intimate partner violence in colombia: Who is at risk? *Social Forces*, 91(2):663–688.

Fruttero, A., Muller, N., and Calvo-Gonzalez, O. (2021). The power and roots of aspirations: A survey of the empirical evidence.

Gächter, S. and Renner, E. (2018). Leaders as role models and 'belief managers' in social dilemmas. *Journal of Economic Behavior & Organization*, 154:321–334.

García-Moreno, C., Hegarty, K., d'Oliveira, A. F. L., Koziol-McLain, J., Colombini, M., and Feder, G. (2015). The health-systems response to violence against women. *The Lancet*, 385(9977):1567–1579.

García-Moreno, C., Pallitto, C., Devries, K., Stöckl, H., Watts, C., and Abrahams, N. (2013). *Global and Regional Estimates of Violence against Women: Prevalence and Health Effects of Intimate Partner Violence and Non-partner Sexual Violence*. World Health Organization.

Gelman, A. and Imbens, G. (2019). Why high-order polynomials should not be used in regression discontinuity designs. *Journal of Business & Economic Statistics*, 37(3):447–456.

Goldberg, I. and Nold, F. C. (1980). Does reporting deter burglars?–an empirical analysis of risk and return in crime. *The Review of Economics and Statistics*, pages 424–431.

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

Google (2022). COVID-19 Community Mobility Reports. Retrieved from https://www.google.com/covid19/mobility/.

Gould, L. C. (1969). The changing structure of property crime in an affluent society. *Social Forces*, 48(1):50–59.

Grace, M. F., Sjoquist, D. L., and Wheeler, L. (2014). The effect of insurance premium taxes on interstate differences in the size of the property-casualty insurance industry. *National Tax Journal*, 67(1):151–181.

Green, J. C. (2013). Indira gandhi: India's destined leader.

Gruber, J. and Lettau, M. (2004). How elastic is the firm's demand for health insurance? *Journal of Public Economics*, 88(7-8):1273–1293.

Gupta, S. K., Uygun, S. V., Ilinich, S., Kumar, R., and Ordukhanyan, E. (2019). Mahatma gandhi: a role model in revolutionary management. *International Journal of Research and Analytical Reviews*, 6(2):993–998.

Harrington, P., Lonsway, K., Moore, M., Smeal, E., and Spillar, K. (2003). Hiring and retaining more women: The advantages to law enforcement agencies. *National Center for Women & Policing a division of the Feminist Majority Foundation. Retrieved from http://www. womenandpolicing. com/pdf/newadvantagesreport. pdf*.

Hasher, L. and Zacks, R. T. (1984). Automatic processing of fundamental information: the case of frequency of occurrence. *American psychologist*, 39(12):1372.

He, Q. and Barkowski, S. (2020). The effect of health insurance on crime: Evidence from the affordable care act medicaid expansion. *Health economics*, 29(3):261–277.

Health Markets (2019). What to Know About Spousal Health Insurance Coverage. Retrieved from https://www.healthmarkets.com/resources/health-insurance/spousal-health-insurance/ Accessed 2 March 2020.

Heath, R. (2014). Women's access to labor market opportunities, control of household resources, and domestic violence: Evidence from bangladesh. *World Development*, 57:32–46.

Helland, E. and Tabarrok, A. (2007). Does three strikes deter? a nonparametric estimation. *Journal of Human resources*, 42(2):309–330.

Hewitt, E. J. (1991). Operational Independence—Myth or Reality? *The Police Journal*, 64(4):321–330.

Hidrobo, M. and Fernald, L. (2013). Cash transfers and domestic violence. *Journal of health economics*, 32(1):304–319.

Hoel, R. (2021). Insurance Facts and Statistics. Retrieved from https://www.bankrate.com/insurance/homeowners-insurance/home-insurance-statistics/ Accessed 12 Jan 2022.

Hyman, A. (1997). Mandatory Reporting of Domestic Violence by Health Care Providers. Retrieved from http://www.futureswithoutviolence.org/userfiles/file/HealthCare/mandatory$_policypaper.pdf$

Hyman, A., Schillinger, D., and Lo, B. (1995). Laws mandating reporting of domestic violence: do they promote patient well-being? *JAMA*, 273(22):1781–1787.

Imbens, G. and Angrist, J. (1994). Identification and estimation of local average treatment effects. *Econometrica*, 62(2):467–475.

Imbens, G. and Kalyanaraman, K. (2012). Optimal bandwidth choice for the regression discontinuity estimator. *The Review of economic studies*, 79(3):933–959.

Imbens, G. W. and Lemieux, T. (2008). Regression discontinuity designs: A guide to practice. *Journal of econometrics*, 142(2):615–635.

India Today (2020). Rajasthan govt releases fresh guidelines for Lockdown. Retrieved from https://www.indiatoday.in/india/story/rajasthan-govt-fresh-guidelines-lockdown-5-coronavirus-covid19-1684002-2020-05-31.

India Today (2021). 82% students of Classes 7 to 10 are fearful of Math: Survey. https://www.indiatoday.in/education-today/latest-studies/story/82-students-classes-7-to-10-are-fearful-of-math-survey-1827619-2021-07-13.

IndiaGov (2021). MyGov - Citizen Participation towards Good Governance. https://www.india.gov.in/spotlight/mygov-citizen-participation-towards-good-governance.

Indianexpress (2015). CRecalling the Emergency years in India. https://indianexpress.com/article/explained/recalling-the-emergency-1975-77-the-emergency-at-work/.

Indianexpress (2018). Rajasthan textbooks to be reviewed, role of Gandhi, Nehru to be restored. https://indianexpress.com/article/india/rajasthan-textbooks-review-gandhi-nehru-congress-5515069/.

Inglehart, R. and Norris, P. (2000). The developmental theory of the gender gap: Women's and men's voting behavior in global perspective. *International Political Science Review*, 21(4):441–463.

Isaac, R. K. and Budryte-Ausiejiene, L. (2015). Interpreting the emotions of visitors: A study of visitor comment books at the grūtas park museum, lithuania. *Scandinavian journal of hospitality and tourism*, 15(4):400–424.

Iyengar, R. (2009). Does the Certainty of Arrest Reduce Domestic Violence? Evidence from Mandatory and Recommended Arrest Laws. *Journal of Public Economics*, 93(1-2):85–98.

Iyer, L., Mani, A., Mishra, P., and Topalova, P. (2012). The Power of Political Voice: Women's Political Representation and Crime in India. *American Economic Journal: Applied Economics*, 4(4):165–93.

Jackson, O. (2020). Punishment and crime: The impact of felony conviction on criminal activity. Technical report, Working Papers.

Jacobs, B. A. and Cherbonneau, M. (2014). Auto theft and restrictive deterrence. *Justice quarterly*, 31(2):344–367.

Jegen, M. and Deats, R. L. (2005). *Mahatma Gandhi, nonviolent liberator: A biography*. New City Press.

Jensen, G. A. and Gabel, J. R. (1992). State mandated benefits and the small firm's decision to offer insurance. *Journal of Regulatory Economics*, 4(4):379–404.

Jensen, R. and Oster, E. (2009). The Power of TV: Cable Television and Women's Status in India. *The Quarterly Journal of Economics*, 124(3):1057–1094.

Jones, T. (2008). The accountability of policing. In *Handbook of policing*, pages 693–724. Willan.

Joseph, G., Javaid, S. U., Andres, L. A., Chellaraj, G., Solotaroff, J. L., and Rajan, S. I. (2017). Underreporting of gender-based violence in kerala, india.

Kahneman, D. (2011). *Thinking, fast and slow*. Macmillan.

Kaiser Family Foundation (2014). The Uninsured: A Primer in America. Retrieved from https://files.kff.org/attachment/the-uninsured-a-primer-key-facts-about-health-insurance-and-the-uninsured-in-america-supplemental-tables.

Kaiser Family Foundation (2019a). How Affordable are 2019 ACA Premiums. Retrieved from https://www.kff.org/health-reform/issue-brief/how-affordable-are-2019-aca-premiums-for-middle-income-people/ Accessed 13 January 2020.

Kaiser Family Foundation (2019b). Marketplace Average Benchmark Premiums. Retrieved from https://www.kff.org/state-category/affordable-care-act/health-insurance-marketplaces/ Accessed 12 January 2020.

Kaiser Family Foundation (2019c). Status of State Action on the Medicaid Expansion Decision. Retrieved from https://www.kff.org/state-category/affordable-care-act/medicaid-and-health-reform-health-reform/ Accessed 2 January 2020.

Kang, J.-A., Hubbard, G. T., and Hong, S. (2019). Gender and credibility in branded storytelling. *Gender in Management: An International Journal*.

Kaufman, I. R. (1980). Criminal procedure in england and the united states: Comparisons in initiating prosecutions. *Fordham L. Rev.*, 49:26.

Kavanaugh, G., Sviatschi, M., and Trako, I. (2018). Women officers, gender violence and human capital: Evidence from women's justice centers in peru.

Kay, R. (2008). Exploring gender differences in computer-related behaviour: Past, present, and future. In *Social information technology: Connecting society and cultural issues*, pages 12–30. IGI Global.

Kay, R. H. (2009). Examining gender differences in attitudes toward interactive classroom communications systems (iccs). *Computers & education*, 52(4):730–740.

Kearney, M. S. and Levine, P. B. (2015). Media influences on social outcomes: The impact of mtv's 16 and pregnant on teen childbearing. *American Economic Review*, 105(12):3597–3632.

Keiser, L. R., Wilkins, V. M., Meier, K. J., and Holland, C. A. (2002). Lipstick and Logarithms: Gender, Institutional Context, and Representative Bureaucracy. *American Political Science Review*, 96(3):553–564.

Kelkar, R., Mokracek, P., Nimbalkar, S. K., and Gandhi, N. (2022). A study of arjuna's qualities and their implications in today's management scenario. *Journal of Positive School Psychology*, 6(11):445–452.

Kelly, L. (2015). *Knowledge and power in prehistoric societies: Orality, memory, and the transmission of culture*. Cambridge University Press.

Kipchumba, E. K., Porter, C., Serra, D., Sulaiman, M., et al. (2021). Infuencing youths' aspirations and gender attitudes through role models: Evidence from somali schools. Technical report.

Kishor, S. and Johnson, K. (2005). Profiling domestic violence: a multi-country study. *Studies in Family Planning*, 36(3):259–261.

Klick, J. and Tabarrok, A. (2005). Using Terror Alert Levels to Estimate the Effect of Police on Crime. *The Journal of Law and Economics*, 48(1):267–279.

Koch, M. T. and Fulton, S. A. (2011). In the defense of women: Gender, office holding, and national security policy in established democracies. *The Journal of politics*, 73(1):1–16.

La Ferrara, E., Chong, A., and Duryea, S. (2012). Soap operas and fertility: Evidence from brazil. *American Economic Journal: Applied Economics*, 4(4):1–31.

Lafortune, J., Riutort, J., and Tessada, J. (2018). Role models or individual consulting: The impact of personalizing micro-entrepreneurship training. *American Economic Journal: Applied Economics*, 10(4):222–45.

Lambert, E. G., Qureshi, H., Hogan, N. L., Klahm, C., Smith, B., and Frank, J. (2015). The Association of Job Variables with Job Involvement, Job Satisfaction, and Organizational Commitment among Indian Police Officers. *International Criminal Justice Review*, 25(2):194–213.

Lauritsen, J. and Catalano, S. (2005). National crime victimization surveys. *Encyclopedia of Social Measurement*, pages 803–808.

Laursen, B. and Faur, S. (2022). What does it mean to be susceptible to influence? a brief primer on peer conformity and developmental changes that affect it. *International Journal of Behavioral Development*, 46(3):222–237.

Lawless, J. L. and Pearson, K. (2008). The primary reason for women's underrepresentation? reevaluating the conventional wisdom. *The Journal of Politics*, 70(1):67–82.

Lee, D. S. (2008). Randomized experiments from non-random selection in us house elections. *Journal of Econometrics*, 142(2):675–697.

Lee, D. S. and Lemieux, T. (2010). Regression discontinuity designs in economics. *Journal of economic literature*, 48(2):281–355.

Lee, D. S. and McCrary, J. (2017). *The deterrence effect of prison: Dynamic theory and evidence*. Emerald Publishing Limited.

Lehman, C. (2022). (Mis)measuring the Shoplifting Crisis. Retrieved from https://thedispatch.com/p/mismeasuring-the-shoplifting-crisis?utm$_source = urlAccessed30January2022.$

Lepage, L.-P. (2020). Are criminals strategic? offender responses to drug sentencing cutoffs. *Labour Economics*, 66:101906.

Levitt, S. D. (1996). The effect of prison population size on crime rates: Evidence from prison overcrowding litigation. *The quarterly journal of economics*, 111(2):319–351.

Levitt, S. D. (1997). Using Electoral Cycles in Police Hiring to Estimate the Effects of Police on Crime. *American Economic Review*, 87(3):270–90.

Levitt, S. D. (1998). Juvenile crime and punishment. *Journal of political Economy*, 106(6):1156–1185.

Lewis, O., Victor, B., et al. (1958). Village life in northern india. studies in a delhi village. *Village Life in Northern India. Studies in a Delhi Village.*

Linek, S. B., Gerjets, P., and Scheiter, K. (2010). The speaker/gender effect: does the speaker's gender matter when presenting auditory text in multimedia messages? *Instructional Science*, 38(5):503–521.

Lizdas, K., Marjavi, A., Durborow, N., and O'Flaherty, A. (2019). Compendium of state and u.s. territory statutes and policies on domestic violence and health care. *Futures Without Violence*, 1(1):1–118.

Lockwood, P. (2006). "someone like me can be successful": Do college students need same-gender role models? *Psychology of women quarterly*, 30(1):36–46.

Lockwood, P. and Kunda, Z. (1997). Superstars and me: Predicting the impact of role models on the self. *Journal of personality and social psychology*, 73(1):91.

Lott, Jr, J. R. and Kenny, L. W. (1999). Did women's suffrage change the size and scope of government? *Journal of political Economy*, 107(6):1163–1198.

MacDonald, Z. (2001). Revisiting the dark figure: A microeconometric analysis of the under-reporting of property crime and its implications. *British Journal of Criminology*, 41(1):127–149.

MacDonald, Z. (2002). Official crime statistics: their use and interpretation. *The Economic Journal*, 112(477):F85–F106.

Maier, S. L. (2008). "i have heard horrible stories..." rape victim advocates' perceptions of the revictimization of rape victims by the police and medical system. *Violence against women*, 14(7):786–808.

Majic, S. (2017). Sending a dear john letter: Public information campaigns and the movement to "end demand" for prostitution in atlanta, ga. *Social Sciences*, 6(4):138.

Malhotra, I. (2014). *Indira Gandhi: A personal and political biography*. Hay House, Inc.

Mansfield, R., Gould, L. C., and Namenwirth, J. Z. (1974). A socioeconomic model for the prediction of societal rates of property theft. *Social Forces*, 52(4):462–472.

Manuel, C. (2000). Mnemonic performances. *The Lancet*, 356(9241):1611.

Marschall, S. (2017). Targeting statues: monument "vandalism" as an expression of sociopolitical protest in south africa. *African Studies Review*, 60(3):203–219.

Martin, S. E. and Jurik, N. C. (2006). *Doing justice, doing gender: Women in legal and criminal justice occupations*. Sage Publications.

Masser, B. M. and Abrams, D. (2004). Reinforcing the Glass Ceiling: The Consequences of Hostile Sexism for Female Managerial Candidates. *Sex Roles*, 51(9-10):609–615.

Mastandrea, S., Fagioli, S., and Biasi, V. (2019). Art and psychological well-being: Linking the brain to the aesthetic emotion. *Frontiers in Psychology*, 10:739.

Matsa, D. A. and Miller, A. R. (2013). A female style in corporate leadership? evidence from quotas. *American Economic Journal: Applied Economics*, 5(3):136–69.

Mawby, R. and Wright, A. (2005). Police Accountability in the United Kingdom. *Human Rights Initiative*, 7:2010.

McCrary, J. (2007). The effect of court-ordered hiring quotas on the composition and quality of police. *American Economic Review*, 97(1):318–353.

McCrary, J. (2008). Manipulation of the running variable in the regression discontinuity design: A density test. *Journal of econometrics*, 142(2):698–714.

McKelway, M. (2019). Experimental evidence on the effects of women's employment. Technical report, Working Paper.

Meier, K., Niessen-Ruenzi, A., and Ruenzi, S. (2020). The impact of role models on women's self-selection in competitive environments. *Available at SSRN 3087862*.

Meier, K. J. and Nicholson-Crotty, J. (2006). Gender, representative bureaucracy, and law enforcement: The case of sexual assault. *Public Administration Review*, 66(6):850–860.

Meier, K. J. and Stewart, J. (1992). Active Representation in Educational Bureaucracies: Policy Impacts. *American Review of Public Administration*, 22(3):157–71.

Miller, A. R. and Segal, C. (2019). Do Female Officers Improve Law Enforcement Quality? Effects on Crime Reporting and Domestic Violence. *The Review of Economic Studies*, 86(5):2220–2247.

Miller, L. S., More, H. W., and Braswell, M. C. (2017). *Effective police supervision*. Routledge.

Ministry of Justice (2013). UK Official Statistics: An Overview of Sexual Offending in England and Wales. Retrieved from https://www.gov.uk/government/statistics/an-overview-of-sexual-offending-in-england-and-wales Accessed 11 Sep 2019.

Morabito, M. S., Pattavina, A., and Williams, L. M. (2017). Active representation and police response to sexual assault complaints. *Journal of crime and justice*, 40(1):20–33.

Morris-Jones, W. (1975). Whose emergency—india's or indira's? *The World Today*, 31(11):451–461.

Mueller-Smith, M. (2015). The criminal and labor market impacts of incarceration. *Unpublished Working Paper*, 18.

Mueller-Smith, M. and Schnepel, K. T. (2021). Diversion in the criminal justice system. *The Review of Economic Studies*, 88(2):883–936.

Muratore, M. G. and Sabbadini, L. L. (2005). Italian survey on violence against women. *Statistical Journal of the United Nations Economic Commission for Europe*, 22(3, 4):265–278.

Myhill, A., Dalgleish, D., Docking, M., Myhill, A., and Yarrow, S. (2003). *The role of police authorities in public engagement*. United Kingdom. Home Office.

Natapoff, A. (2011). Misdemeanors. *S. Cal. L. Rev.*, 85:1313.

Natapoff, A. (2015). Misdemeanors. *Annual Review of Law and Social Science*, 11:255–267.

National Association of Insurance Commissioners (2017). A Guide to State Retaliatory Taxes, Fees, Deposits and Other Requirements. Retrieved from https://www.naic.org/documents/prod$_{s}erv_{l}egal_{r}et_{z}u.pdf Accessed 2 July 2020.$

National Association of Insurance Commissioners (2019). Home Insurance Facts and Statistics. Retrieved from https://content.naic.org/sites/default/files/publication-hmr-zu-homeowners-report.pdf Accessed 12 March 2022.

National Women's Law Center (2022). Resilient But Not Recovered. Retrieved from https://nwlc.org/wp-content/uploads/2022/03/FINAL-NWLC-Resilient-But-Not-Recovered-3.29.22.pdf.

Neggers, Y., Goldenberg, R., Cliver, S., and Hauth, J. (2004). Effects of domestic violence on preterm birth and low birth weight. *Acta obstetricia et gynecologica Scandinavica*, 83(5):455–460.

Njuki, R., Okal, J., Warren, C. E., Obare, F., Abuya, T., Kanya, L., Undie, C.-C., Bellows, B., and Askew, I. (2012). Exploring the effectiveness of the output-based aid voucher program to increase uptake of gender-based violence recovery services in kenya: a qualitative evaluation. *BMC Public Health*, 12(1):1–8.

Office of Management and Budget (2010). 2010 Office of Management and Budget (OMB) Standards. Retrieved from https://www.census.gov/programs-surveys/metro-micro/about.html Accessed 11 Sep 2020.

Oram, S., Khalifeh, H., and Howard, L. M. (2017). Violence against Women and Mental Health. *The Lancet Psychiatry*, 4(2):159–170.

Overstreet, N. M. and Quinn, D. M. (2013). The intimate partner violence stigmatization model and barriers to help seeking. *Basic and applied social psychology*, 35(1):109–122.

Owens, E. G. (2009). More time, less crime? estimating the incapacitative effect of sentence enhancements. *The Journal of Law and Economics*, 52(3):551–579.

Özler, B., Çelik, Ç., Cunningham, S., Cuevas, P. F., and Parisotto, L. (2021). Children on the move: Progressive redistribution of humanitarian cash transfers among refugees. *Journal of Development Economics*, 153:102733.

Palermo, T., Bleck, J., and Peterman, A. (2014). Tip of the iceberg: reporting and gender-based violence in developing countries. *American journal of epidemiology*, 179(5):602–612.

Pande, R. (2003). Can Mandated Political Representation Increase Policy Influence for Disadvantaged Minorities? Theory and Evidence from India. *American Economic Review*, 93(4):1132–1151.

Panyamanee, P. (2022). Bhimrao ramji ambedkar's image and thought as perceived in thailand from 1975 to 2017. *Contemporary Voice of Dalit*, page 2455328X211066198.

Perera, L. C. R. and Chaminda, J. W. D. (2013). Corporate social responsibility and product evaluation: The moderating role of brand familiarity. *Corporate Social Responsibility and Environmental Management*, 20(4):245–256.

Perova, E. and Reynolds, S. A. (2017). Women's Police Stations and Intimate Partner Violence: Evidence from Brazil. *Social Science & Medicine*, 174:188–196.

Peterson, C., DeGue, S., Florence, C., and Lokey, C. N. (2017). Lifetime economic burden of rape among us adults. *American journal of preventive medicine*, 52(6):691–701.

Peterson, C., Xu, L., and Florence, C. (2021). Average medical cost of fatal and non-fatal injuries by type in the usa. *Injury prevention*, 27(1):24–33.

Pinotti, P. (2017). Clicking on heaven's door: The effect of immigrant legalization on crime. *American Economic Review*, 107(1):138–68.

Plichta, S. B. and Falik, M. (2001). Prevalence of Violence and Its Implications for Women's Health. *Women's Health Issues*, 11(3):244–258.

Pollak, R. A. (2004). An intergenerational model of domestic violence. *Journal of Population Economics*, 17(2):311–329.

Porter, C. and Serra, D. (2020). Gender differences in the choice of major: The importance of female role models. *American Economic Journal: Applied Economics*, 12(3):226–54.

Powley, E. (2007). Rwanda: The impact of women legislators on policy outcomes affecting children and families. *Background Paper, State of the World's Children*.

Prakash, G. (2019). *Emergency Chronicles: Indira Gandhi and Democracy's Turning Point*. Princeton University Press.

Prayag, G., Hosany, S., and Odeh, K. (2013). The role of tourists' emotional experiences and satisfaction in understanding behavioral intentions. *Journal of Destination Marketing & Management*, 2(2):118–127.

Price, S. (2020). Average Cost of Health Insurance. Retrieved from https://www.valuepenguin.com/average-cost-of-health-insurance Accessed 12 August 2020.

Rackimuthu, S., Hasan, M. M., Bardhan, M., and Essar, M. Y. (2022). Covid-19 vaccination strategies and policies in india: the need for further re-evaluation is a pressing priority. *The International Journal of Health Planning and Management*, 37(3):1847–1850.

Rahman, M., Hoque, M. A., and Makinoda, S. (2011). Intimate partner violence against women: Is women empowerment a reducing factor? a study from a national bangladeshi sample. *Journal of Family Violence*, 26(5):411–420.

Rajaraman, I. and Gupta, M. (2012). Public expenditure choices and gender quotas. *Indian Growth and development review*.

Ranji, U., Salganicoff, A., Rousseau, D., et al. (2019). Barriers to care experienced by women in the united states. *Jama*, 321(22):2154–2154.

Ray, D. (2006). Aspirations, poverty, and economic change. *Understanding poverty*, 1:409–421.

Regan, P. M. and Paskeviciute, A. (2003). Women's access to politics and peaceful states. *Journal of Peace Research*, 40(3):287–302.

Riccucci, N. M., Van Ryzin, G. G., and Lavena, C. F. (2014). Representative bureaucracy in policing: Does it increase perceived legitimacy? *Journal of public administration research and theory*, 24(3):537–551.

Rice, B. (2017). A reform long overdue: Raising virginia's felony grand larceny threshold. *Rich. Pub. Int. L. Rev.*, 21:1.

Riley, E. (2022). Role models in movies: the impact of queen of katwe on students' educational attainment. *The Review of Economics and Statistics*, pages 1–48.

Robertson, R., Stremikis, K., Collins, S. R., Doty, M. M., and Davis, K. (2012). Jobs without benefits: the health insurance crisis faced by small businesses and their workers. *Issue brief (Commonwealth Fund)*, 28:1–24.

Sakalli-Ugurlu, N. and Beydogan, B. (2002). Turkish College Students' Attitudes toward Women Managers: The Effects of Patriarchy, Sexism, and Gender Differences. *The Journal of psychology*, 136(6):647–656.

Sampathkumar, M. (2015). Bhimrao Ramji Ambedkar Early Life A Study. https://citeseerx.ist.psu.edu/document?repid=rep1type=pdfdoi=26b3d6d62048f588d7bb30a

Sampson, F. (2012). Hail to the chief?—how far does the introduction of elected police commissioners herald a us-style politicization of policing for the uk? *Policing: a journal of policy and practice*, 6(1):4–15.

Schechter, L. (2021). Do shelters reduce domestic violence?

Scrase, T. J. (2002). Television, the middle classes and the transformation of cultural identities in west bengal, india. *Gazette (Leiden, Netherlands)*, 64(4):323–342.

Sekhri, S. and Storeygard, A. (2014). Dowry Deaths: Response to Weather Variability in India. *Journal of Development Economics*, 111:212–223.

Selden, S. C. (1997). *The promise of representative bureaucracy: Diversity and responsiveness in a government agency*. ME Sharpe.

Short, P. F. (1998). Gaps and transitions in health insurance: what are the concerns of women? *Journal of Women's Health*, 7(6):725–737.

Siegel, R. B. (1995). The rule of love: Wife beating as prerogative and privacy. *Yale LJ*, 105:2117.

Sims, K. M. (2021). Seeking safe harbors: Emergency domestic violence shelters and family violence. Technical report, Working Paper.

Skogan, W. G. (1975). Measurement problems in official and survey crime rates. *Journal of criminal justice*, 3(1):17–31.

Skogan, W. G. (1976). Citizen Reporting of Crime Some National Panel Data. *Criminology*, 13(4):535–549.

Snyder, C. R. (2002). Hope theory: Rainbows in the mind. *Psychological inquiry*, 13(4):249–275.

Soares, R. R. (2004a). Crime reporting as a measure of institutional development. *Economic Development and cultural change*, 52(4):851–871.

Soares, R. R. (2004b). Development, crime and punishment: accounting for the international differences in crime rates. *Journal of development Economics*, 73(1):155–184.

Sobkoviak, R. M., Yount, K. M., and Halim, N. (2012). Domestic violence and child nutrition in liberia. *Social science & medicine*, 74(2):103–111.

Soni, A. and Kumari, S. (2017). The role of parental math anxiety and math attitude in their children's math achievement. *International Journal of Science and Mathematics Education*, 15(2):331–347.

Soroka, S., Gidengil, E., Fournier, P., and Nir, L. (2016). Do women and men respond differently to negative news? *Politics & Gender*, 12(2):344–368.

Soroka, S. N. (2014). *Negativity in democratic politics: Causes and consequences*. Cambridge University Press.

Stevens, J. S. and Hamann, S. (2012). Sex differences in brain activation to emotional stimuli: a meta-analysis of neuroimaging studies. *Neuropsychologia*, 50(7):1578–1593.

Stöckli, S. and Hofer, D. (2020). Susceptibility to social influence predicts behavior on facebook. *PloS one*, 15(3):e0229337.

Stout, J. G., Dasgupta, N., Hunsinger, M., and McManus, M. A. (2011). Steming the tide: using ingroup experts to inoculate women's self-concept in science, technology, engineering, and mathematics (stem). *Journal of personality and social psychology*, 100(2):255.

Suarez, E. and Gadalla, T. M. (2010). Stop blaming the victim: A meta-analysis on rape myths. *Journal of interpersonal violence*, 25(11):2010–2035.

Sun, L. and Abraham, S. (2021). Estimating dynamic treatment effects in event studies with heterogeneous treatment effects. *Journal of Econometrics*, 225(2):175–199.

The Hindu (2020a). Rajasthan announces complete lockdown till March 31. Retrieved from https://www.thehindu.com/news/national/other-states/coronavirus-rajasthan-announces-complete-lockdown-till-march-31/article31133432.ece.

The Hindu (2020b). Rajasthan closes schools, cinema halls till March 30. Retrieved from https://www.thehindu.com/news/national/other-states/coronavirus-rajasthan-closes-schools-cinema-halls-till-march-30/article31067344.ece.

The Hindu (2021). Rajasthan modifies lockdown to ease restrictions. https://www.thehindu.com/news/national/other-states/rajasthan-modifies-lockdown-to-ease-restrictions/article34757011.ece.

Thielemann, G. S. and Stewart Jr, J. (1996). A Demand-side Perspective on the Importance of Representative Bureaucracy: AIDS, Ethnicity, Gender, and Sexual Orientation. *Public Administration Review*, pages 168–173.

Times of India (2014). Phobia keeps students away from math: Study. https://timesofindia.indiatimes.com/education/news/phobia-keeps-students-away-from-math-study/articleshow/45652092.cms.

Tol, W., Murray, S., Lund, C., Bolton, P., Murray, L., Davies, T., Haushofer, J., Orkin, K., Witte, M., Salama, L., et al. (2019). Can mental health treatments help prevent or reduce intimate partner violence in low-and middle-income countries? a systematic review. *BMC women's health*, 19(1):1–15.

Trusts, P. C. (2017). The effects of changing felony theft thresholds. *Chartbook, April. Philadelphia, PA: The Pew Charitable Trusts*.

Tung, V. W. S. and Ritchie, J. B. (2011). Exploring the essence of memorable tourism experiences. *Annals of tourism research*, 38(4):1367–1386.

Tur-Prats, A. (2019). Family types and intimate partner violence: A historical perspective. *Review of Economics and Statistics*, 101(5):878–891.

Ulrich, Y. C., Cain, K. C., Sugg, N. K., Rivara, F. P., Rubanowice, D. M., and Thompson, R. S. (2003). Medical Care Utilization Patterns in Women with Diagnosed Domestic Violence. *American Journal of Preventive Medicine*, 24(1):9–15.

United Nations (2012). Women in politics the fight to end violence against women. Retrieved from https://www.un.org/en/chronicle/article/women-politics-fight-end-violence-against-women Accessed 12 May 2020.

United Nations Office on Drugs and Crime (2018). Global Report on Trafficking in Persons 2018. Retrieved from https://www.unodc.org/documents/data-and-analysis/glotip/2018/GLOTiP$_{2}018_{B}OOK_{w}eb_{s}mall.pdf Accessed 15 Oct 2020.$

Ward, B. (2020). The impact of medicaid expansion on states' budgets. *Commonwealth Fund, May*, 5.

Watts, D. J. and Dodds, P. S. (2007). Influentials, networks, and public opinion formation. *Journal of consumer research*, 34(4):441–458.

Weissman, M. M. and Olfson, M. (1995). Depression in women: implications for health care research. *Science*, 269(5225):799–801.

Wen, H., Hockenberry, J. M., and Cummings, J. R. (2017). The effect of medicaid expansion on crime reduction: Evidence from hifa-waiver expansions. *Journal of Public Economics*, 154:67–94.

Wentz, E. and Archbold, C. A. (2012). Police perceptions of sexual assault victims: Exploring the intra-female gender hostility thesis. *Police Quarterly*, 15(1):25–44.

West, J. (2019). Learning by doing in law enforcement. Technical report, Working paper.-. 05.

White House (2021). National Strategy on Gender Equity and Equality. Retrieved from https://www.whitehouse.gov/wp-content/uploads/2021/10/National-Strategy-on-Gender-Equity-and-Equality.pdf.

Whitley Jr, B. E. (1997). Gender differences in computer-related attitudes and behavior: A meta-analysis. *Computers in human behavior*, 13(1):1–22.

Williams, L. M., Barton, M. J., Kemp, A. H., Liddell, B. J., Peduto, A., Gordon, E., and Bryant, R. A. (2005). Distinct amygdala–autonomic arousal profiles in response to fear signals in healthy males and females. *Neuroimage*, 28(3):618–626.

Wilson, W. J. (2012). *The truly disadvantaged: The inner city, the underclass, and public policy*. University of Chicago Press.

Wolf, M. E., Ly, U., Hobart, M. A., and Kernic, M. A. (2003). Barriers to seeking police help for intimate partner violence. *Journal of family Violence*, 18(2):121–129.

World Economic Forum (2020). Global gender gap report 2020.

World Health Organization (2005). *WHO multi-country study on women's health and domestic violence against women: Initial results on prevalence, health outcomes and women's responses*. World Health Organization.

WorldBank (2014). *World development report 2015: Mind, society, and behavior*. The World Bank.

Zelliot, E. (2008). Understanding dr. br ambedkar. *Religion Compass*, 2(5):804–818.

Zizzo, D. J. (2010). Experimenter demand effects in economic experiments. *Experimental Economics*, 13(1):75–98.