Safe Travels: Transport Development and Women's Safety in India*

Anisha Garg†

August 28, 2024

Abstract

Beyond fostering inclusion and connectivity, does investing in mass transit impact crimes against women? I study the impact of the expansion of the metro rail system in Delhi, India, on reported crimes against women. Preparing a novel data set of the universe of crime reports at the police station-day level, and leveraging the staggered opening of metro stations across these police jurisdictions between 2016 and 2019, I find a 29 percent decrease in reported incidents of sexual harassment against women in public spaces following the opening of the first metro station in the police jurisdiction. This reduction in reported sexual harassment is neither linked to an overall decrease in reported non-gender-specific crimes in public spaces, nor substituted by an increase in such crimes in metro stations or trains. Further analysis of crime reports from alternative public transport networks, combined with area safety data from a crowd-sourced mobile application, suggests that this reduction is driven by the enhanced safety provided by the metro network.

Keywords: Crimes against women, state transit networks, Delhi

JEL Codes: H72, J16, K42, O18, R41

*I am grateful to my supervisors, Robert Akerlof, Debraj Ray and Andreas Stegmann, for their continuous guidance and support. I also thank Wiji Arulampalam, Yatish Arya, Manuel Bagues, Dan Bernhardt, Sonia Bhalotra, Nathan Canen, Nikhil Datta, Mirko Draca, Bhaskar Dutta, Ludovicca Gazze, Bishnupriya Gupta, Clément Imbert, Amrita Kulka and the audience at the Warwick CAGE-AMES workshop in 2024 for their helpful comments. I am also grateful to the Joint Commissioner of Delhi Police for helping me understand the nuances of police jurisdictions and crime data. I would like to thank Kalpana Viswanath and the team at SafetiPin for sharing the SafetiPin mobile app data.

†The University of Warwick, Department of Economics (e-mail: anisha.garg@warwick.ac.uk)

1 Introduction

Developing economies invest in the advancement of transportation infrastructure to foster social inclusion and promote connectivity across different regions within the country.[1] However, these well-intentioned investments can give rise to complex implications in the realm of women's safety, a pressing global concern. As underscored by the World Health Organization (WHO), a substantial 31% of women aged 15-49 years in 2018 were victims of crimes perpetrated by unfamiliar individuals or close relatives at least once during their lifetimes. The gravity of this issue intensifies in low and middle-income countries, where this percentage increases to 34-36%.[2] These crimes not only inflict immediate harm but also create an atmosphere of fear and insecurity, impeding women's engagement in education and employment.[3] Thus, it becomes imperative to delve into the intricate dynamics surrounding the expansion of transport infrastructure within lower-middle-income countries and its potential influence on crimes against women.

Investment in transport infrastructure may impact reported crimes against women through multiple potential channels. Improved surveillance, well-lit stations, and commuting options closer to home may minimise time spent traversing dimly lit areas, leading to fewer reported crimes (*safety* effect). Conversely, increased mobility resulting from improved transportation access might expand the number of potential victims and opportunities for crime, potentially leading to more reported incidents (*volume* effect). Furthermore, this expanded mobility could reconfigure the geographic distribution of crimes as people traverse diverse locales to access new transportation options (*spatial* effect). Additionally, higher reported crimes might not only indicate actual increases but also reflect greater awareness and empowerment of women to report offenses, facilitated by access to economic opportunities (*reporting* ef-

[1](Dominguez Gonzalez et al., 2020); `https://www.eib.org/en/stories/developing-countries-transport-infrastructure`

[2]`https://www.who.int/publications/i/item/9789240022256`

[3]For example, Borker (2021) demonstrate that women may have to compromise their educational choices by selecting institutions based on the safety of travel routes. Additionally, Siddique (2022) illustrate how media-reported sexual assaults can negatively affect women's participation in the labor force.

fect). Hence, a comprehensive understanding of these multifaceted effects and their interplay is vital for an exhaustive analysis of the impact of transport expansion on reported crimes against women.

I study these issues in the context of Delhi, India, which has made considerable progress in improving transport infrastructure with the expansion of the metro rail system in recent years.[4] At the same time, it remains quite behind in providing women with safe and unimpeded access to public facilities.[5] In particular, I ask the following question, does the opening of first metro station in a police jurisdiction impact reported crimes against women? I exploit the exogenous variation in the staggered opening of new metro stations in Delhi to causally establish their impact on reported crimes against women at the police jurisdiction level. I employ two primary data sets for this analysis. First, I construct a novel fine-grained data set of the universe of reported criminal incidents across Delhi from 2016 to 2019, discerned at the level of police station and day.[6] Second, I utilize data about the opening dates and geographical locations of metro stations, thereby identifying the treated jurisdictions.[7]

To estimate the impact of the opening of the first metro station on reported crimes against women, I consider the recent advancements in the Difference-in-Differences (DiD) methodology and adopt the estimation approach in Callaway & Sant'Anna (2021).[8] Using Callaway & Sant'Anna (2021), I compare jurisdictions where a metro station opens with those without a metro station until 2019 ('never treated') and those where a station has not opened until that time but will eventually open in the analysis period ('not-yet-treated').

[4]Delhi is the national capital of India.

[5](Viswanath & Mehrotra, 2007); https://edition.cnn.com/travel/article/worst-transport-for-women/index.html; https://timesofindia.indiatimes.com/city/delhi/delhi-government-took-slew-of-steps-to-ensure-womens-safety/articleshow/74732463.cms

[6]Appendix Section A.1 elucidates the data preparation strategy.

[7]Figure 1 in Section 4 and Table A2 in Appendix Section A.2 show the treated police jurisdictions and the opening dates of metro lines, respectively.

[8]The recent literature highlights the challenges posed by using Two-Way Fixed Effects (TWFE) regression in the presence of staggered treatment and heterogeneous treatment effects. To understand the extent of this issue in my context, I perform the diagnostic test in Goodman-Bacon (2021), which reveals that around 4% of the TWFE estimates are influenced by the 'forbidden' comparison between later treated jurisdictions and those treated earlier as the control. I discuss the rationale for adopting the estimation strategy provided in Callaway & Sant'Anna (2021) over other alternatives (Sun & Abraham, 2021; Borusyak et al., 2021; de Chaisemartin & D'Haultfœuille, 2020) in Sub-section 5.2.

My initial findings indicate that reported crimes related to sexual harassment in public spaces ('Harassment on Street' crimes or 'HoS' crimes going forward) against women per 100,000 population decrease by 29.34% in the treated jurisdictions compared to the control jurisdictions in the quarters following the opening of the first metro station. This decrease is marginally significant, with a p-value of 0.077.

Since metro stations have separate police units ('metro police stations' going forward) responsible for incidents within their premises and on metro rail, there is a possibility that reported HoS crimes within the purview of metro police stations substitute those within the jurisdiction of general police stations.[9] So, I examine the crime reports from metro police stations with a new metro station between 2016 and 2019. A total of 42 such reports at the metro police station compared to a decrease of 271 such reports (overall treatment effect of the opening of metro stations) at general police stations suggests that the substitution of reported HoS crimes between general police stations and metro police stations might not be occurring. It is also possible that this decrease in reported HoS crimes against women is part of a broader decline in street crimes due to heightened security near metro stations. However, I do not find any significant impact of the opening of the first metro station on non-gender-specific reported crimes in public spaces ('Non-gender-specific street crimes' going forward). Furthermore, there is no significant impact of the treatment on the overall reported crimes against women or reported domestic violence crimes.

Next, I validate the underlying identifying assumption with the help of event study plots and demonstrate the plausibility of parallel trends in the evolution of crime statistics between the treated and the control jurisdictions (Callaway & Sant'Anna, 2021). These results remain robust to comparison with an alternate control group ('never treated' jurisdictions) and to the application of alternative estimators proposed in Sun & Abraham (2021), Borusyak et al. (2021), and de Chaisemartin & D'Haultfœuille (2020).

Finally, I present a preliminary analysis of potential mechanisms and delineate the next

[9] For further details, see Section 3.

steps to disentangle them further. I provide three pieces of evidence supporting the *safety* effect. First, I analyze reported HoS crimes against women on buses, an alternative mass public transport system, which are identified using text analysis on the FIR description. The Delhi Metro organizes stations by lines, and between 2016 and 2019, new stations opened on the magenta, pink, and violet lines. The magenta line, in particular, reduced travel time between South-West and South-East Delhi by nearly half. This is visible in the surge in metro ridership post opening of stations on the magenta line and a more pronounced reduction in reported HoS crimes against women in jurisdictions where the first station opened on the magenta line. Since buses serve as a substitute for metros, I also see a reduction in reported HoS crimes perpetrated on women in buses or bus stops in the jurisdictions where stations on the magenta line opened.[10] Second, using area safety data from SafetiPin, I find that areas within 500 metres of the metro station receive a higher safety score after the opening of the metro station compared to areas within 500 and 1,000 metres.[11] Finally, I reinforce the safety of metro stations using survey data from the International Centre for Research on Women and Observer Research Foundation. Collectively, these findings strongly support the *safety* effect. Next, given the decrease in reported HoS crimes against women in jurisdictions where the first metro station opened, along with the increase in metro ridership, the *safety* effect may outweigh any potential *volume* effect. Lastly, I discuss the next steps I plan to take to disentangle these effects further.

1.1 Related Literature

I contribute to the existing literature on the economics of crime, particularly focusing on those against women (Card & Dahl, 2011; Bhuller et al., 2013; Miller & Segal, 2018). Specifically, I contribute to the economic literature on the determinants of reported crimes against women in India. For instance, higher female representation in local government councils led to higher

[10] I will supplement this analysis with the bus ridership to show the substitution between the buses and the metros.

[11] For further details, see Section 6.

reported crimes against women (Iyer et al., 2012), and an alcohol ban in an Indian state led to less violence against women (Luca et al., 2015). Jassal (2020); Amaral et al. (2021) show that opening up of women police stations affected reported crimes against women.[12] To the best of my knowledge, this paper is the first to study how expanding transport facilities affects reported crimes against women in India, with the results highlighting the importance of providing safe and easy access to transport facilities to women. Moreover, I create and utilize a novel micro-level data set at the granularity of police stations, something which most previous studies studying crime in India do not.[13]

Additionally, I contribute to the literature on how mass transport and transit infrastructure affect regional crime dynamics (Billings et al., 2011; Khanna et al., Working Paper). While prior research has explored the relationship between transport expansion and overall crime rates, I delve into a more nuanced exploration to understand the intricate relationship between mass transport development and women's safety, a previously unexplored topic.

The rest of the proposal is structured as follows. Section 2 elucidates the conceptual framework. Section 3 discusses the background and the data sources. Section 4 outlines the estimation strategy. Section 5 discusses the initial findings and robustness checks. Section 6 provides a brief overview of the preliminary analysis and outlines potential steps to be undertaken to disentangle the various mechanisms. Finally, Section 7 concludes briefly.

2 Conceptual Framework

I analyse the ramifications of public transport infrastructure expansion in Delhi on reported crimes against women. There are several potential channels through which this impact may manifest as discussed below.

On the one hand, we might expect to see a decrease in reported crimes against women

[12] There are others like McDougal et al. (2021) which analyse the impact of the notorious '*Nirbhaya*' gang rape case of December 2012 in Delhi on rape reporting rates. See `https://www.bbc.com/news/world-asia-35115974` for information on the case.

[13] Jassal (2020) uses police station-level data for the state of Haryana.

through the 'safety' effect. Well-illuminated metro transit stations and the metro rail, along with security personnel and surveillance cameras during operational hours, offer a secure public transport option. Moreover, when a new metro station is constructed, it also improves lighting and security in the surrounding areas, which may deter potential criminals, instilling a sense of security that encourages women and their families to use public transport more confidently.

Conversely, expanding transport facilities could inadvertently increase reported crimes through the 'volume' effect. With improved access to different areas of a city, women might switch to public transport for commuting purposes or increase their travel frequency. For instance, women might choose to pursue educational and employment opportunities or engage in leisure activities. This increase in travel frequency might elevate their exposure to potential criminal activities.

Moreover, the impact of the introduction of new transport facilities might extend well beyond its immediate borders, exerting influence on neighbouring areas. The areas with these new transit stations might attract females from neighbouring areas to access safe public transport. However, this influx of females in these areas might lead to a rise in overall criminal incidents. This geographical relocation of crimes to areas with transit stations can be termed a 'spatial' effect.

Furthermore, the relationship between reported crimes and actual incidents might diverge due to the 'reporting' effect. As women gain greater exposure and independence with access to education and employment opportunities, their tolerance of domestic violence, harassment, or other atrocities might reduce, leading to a greater willingness to report such incidents. This could contribute to an apparent increase in reported crimes, even if the actual prevalence declines.

To empirically explore the multifaceted relationship between transport infrastructure expansion and reported crimes against women, I leverage the staggered opening of the first metro station within a police jurisdiction in Delhi between 2016 and 2019. I aim to disentangle the specific channels underlying any observed effects, specifically the safety, volume, spatial, and reporting effects. The potential data and methods for untangling these mechanisms are discussed in Section 6.

3 Background and Data

3.1 Reported Crimes against Women

Crimes against women encompass a broad spectrum, ranging from theft and kidnapping to the most severe offenses, such as sexual harassment, domestic violence, rape, and murder. In Delhi, the national capital territory of India, the Delhi Police, overseen by the Ministry of Home Affairs, Government of India, plays a crucial role in crime prevention, detection, and the maintenance of law and order. The city is demarcated into 209 discrete police jurisdictions, each responsible for addressing incidents within its defined boundaries.[14] These boundaries have evolved over time based on crime rates, leading me to consolidate them into 147 jurisdictions for this analysis to ensure consistency between 2016 and 2019. Apart from these 147 general police stations, 16 separate metro police stations have been set up to deter and investigate crimes occurring on the metro rail or within the metro station.[15]

Every incident is reported to the police station under whose jurisdiction it occurred, where a comprehensive first-hand account of events and circumstances is captured in a First Information Report (FIR).[16] Central to the investigative process, these FIRs from all the

[14] `https://delhipolice.gov.in/doc/Citizen_Charter.pdf` At present, the sanctioned strength of Delhi police is 83,762, spread across 6 ranges, 15 districts, and 209 police stations, making it the largest metropolitan police in the world, larger than London, Paris, New York, and Tokyo. For further information, see `https://delhipolice.gov.in/history`.

[15] `https://www.delhimetrorail.com/list-of-metro-police-stations`

[16] The on-duty police officer presents the FIR to the local magistrate, within whose geographical jurisdiction the crime transpired (`https://delhipolice.gov.in/doc/Citizen_Charter.pdf`). Additionally, Section 177 of The Criminal Procedure Code, 1973 stipulates that every crime committed must undergo investigation and subsequent trial within the precincts of the local court vested with authority over the area in which the crime was perpetrated (`https://legislative.gov.in/sites/default/files/A1974-02.pdf`). For further details on how to file an FIR, refer to `https://legalbots.in/blog/how-to-file-a-complaint-against-domestic-violence-in-india`.

police stations have been meticulously scanned and cataloged on the Delhi Police's website since 2011.[17] Each FIR serves as a comprehensive dossier, outlining specific details of the crime: the legal sections and statutes underpinning its registration, intricate details of the informant, the accused, and the reporting officer, a thorough statement from the informant or victim recounting the crime, an itemized inventory of involved property, potential reporting delays, a precise description of the incident's location, and exact date and time stamps for both the occurrence and its reporting. Additionally, it outlines ongoing actions and provisions for transferring the case to alternative police stations.

I create a fine-grained data set of approximately 300,000 distinct crime reports spanning 2016 to 2019, forming the foundation for calculating the analysis's outcome variable. These scanned PDF documents were extracted from the Delhi Police website and parsed into a machine-readable format amenable to analytical exploration using *Optical Character Recognition* (OCR) techniques in Python. Presently, the deployment of *Named Entity Recognition* (NER) in Python is in progress, aimed at adeptly identifying and extracting names of victims and accused parties embedded within the crime narratives. A synthesis of factors, including designated legal 'Section(s),' the capabilities of the *NamSor* package in R, and gender indicators such as 'W/O' (wife of), 'D/O' (daughter of), and 'S/O' (son of), will collectively illuminate instances where crimes specifically target women. Currently, I have focused on using the 'Section(s)' to identify crimes against women. Appendix Section A.1 provides a sample FIR, elucidates the employed data preparation strategy, and lists the different sections of the Indian Penal Code 1860 selected to create the respective crime categories.

This divergence from the conventional reliance on crime statistics, as disseminated by the National Crime Records Bureau (NCRB) in existing literature, is motivated by the need for granular crime data at the police station level—a level of detail absent in NCRB's district-level aggregates.[18] Furthermore, the NCRB data exhibits certain limitations: it does

[17] `http://59.180.234.21:8080/citizen/firSearch.htm`

[18] For further details on NCRB, see `https://ncrb.gov.in/en`.

not encompass reports retracted due to false accusations or victim withdrawals influenced by familial or societal pressures. Additionally, the aggregated framework may inadvertently obscure the gravity of certain heinous crime categories grouped under broader sections of the Indian Penal Code (e.g., murder and rape), potentially resulting in an under-representation of certain crime typologies. Lastly, NCRB statistics are published at an annual level, while FIRs offer precise reporting dates for each crime, enhancing their temporal accuracy.[19]

3.2 Delhi Metro

The inauguration of Delhi Metro's first line on December 25^{th}, 2002, marked a significant milestone in the urban transportation landscape of the city. A collaborative effort between the Government of India and the Delhi state government, the Delhi Metro Rail Corporation, came into being to spearhead this transformative initiative. Since its inception in 2002, the Delhi Metro has unfurled an intricate lattice of 12 lines that traverse 286 stations, seamlessly linking diverse locales in Delhi, Gurugram, and Noida.[20] The evolution of this network has unfolded through a series of carefully planned phases, with three phases currently operational and a fourth phase making progressive strides. A comprehensive summary of each phase is presented in Table A2 in Appendix Section A.2, complemented by the visual representation of the current metro network in Delhi in Figure A7. The cumulative expanse of this network now spans approximately 243 miles.[21]

In the context of this analysis, I will focus on the 24 new metro stations opened between 2016 and 2019 that serve as the first stations within their respective police jurisdictions.[22] This strategic selection forms the cornerstone of an event-study analysis, poised to unveil insights into the extensive margin impact of newly opened metro stations on reported incidents

[19] For a detailed understanding of limitations of NCRB data, see `https://ncrb.gov.in/sites/default/files/CII%20Disclaimer%202020.pdf`.

[20] Gurugram and Noida lie on the boundary of Delhi and belong to the states of Haryana on the North and Uttar Pradesh in the East, respectively.

[21] For further details, visit `https://www.delhimetrorail.com/pages/en/introduction`.

[22] The opening dates are sourced from `https://www.delhimetrorail.com/pages/en/present-network`, while the geographic parameters including latitude, longitude and the associated police jurisdiction are extracted from `https://www.delhimetrorail.com`.

of crimes against women.

3.3 Population

The population data is sourced from the 2011 census, which is organized at the level of wards under Municipal Corporation of Delhi (MCD), Delhi Cantonment and North Delhi Municipal Corporation (NDMC).[23] To seamlessly integrate this data, I superimposed these ward boundaries onto the police jurisdictions. By identifying the shared intersection between these boundaries, I derived the shared regions and subsequently computed the population for each jurisdiction. This population information is instrumental in computing per capita reported crime rates.

4 Empirical Strategy

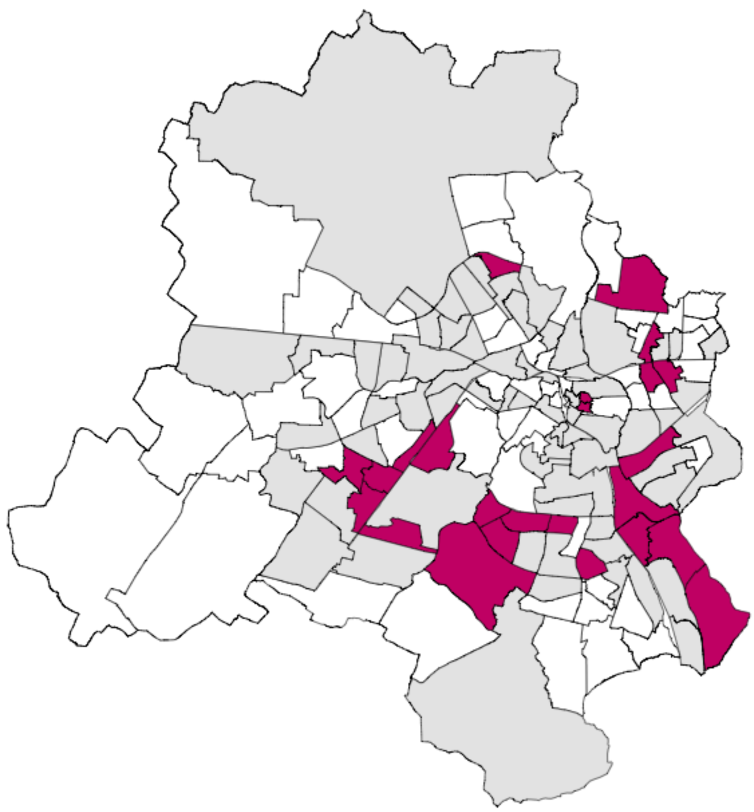
The Delhi metro system's phased construction and operation of stations provide exogenous variation. This variation stems from the land investigations and speed of getting the required approvals. For instance, the construction work on all the stations on the Pink line of the Delhi metro commenced in 2011. However, due to land-related challenges, these stations were opened in 4 different quarters between 2016 and 2019.[24] So, to causally establish the impact of the opening of the first metro station within a police jurisdiction on reported crimes against women, I leverage staggered opening dates of 24 metro stations between 2016 and 2019 for an event study analysis.[25]

The analysis will be at the police jurisdiction level, justified by the necessity to report crimes to the police station under whose purview it occurs. So, given the staggered opening of the metro stations, the jurisdictions can be divided into three groups. First is the treatment group comprising 24 jurisdictions where the first metro station opened during the study

[23] `https://censusindia.gov.in/census.website/data/census-tables`
[24] `https://sundayguardianlive.com/news/metro-complete-phase-3-year`
[25] For details on opening dates and routes covered, see Table A2 and Figure A7 in Appendix Section A.2.

period. Second is the never-treated group, consisting of 55 jurisdictions without a metro station by 2019.[26] Third, the already treated group, encompassing 68 jurisdictions with a metro station before 2016, is omitted from the analysis. Once treated, a jurisdiction remains as such, given that a constructed metro station is not dismantled consciously. Thus, treatment is an absorbing state. Figure 1 and Table 1 illustrates the three groups and the quarterly count of treated jurisdictions, respectively.

Figure 1: Treated Police Jurisdictions



Notes: Treated: 24 police jurisdictions in pink are the ones where the first metro station opened in the treatment period between 2016 and 2019. *Never treated*: 55 in white are the ones without a metro station till 2019. *Already treated*: 68 in grey are those with the first metro station built before 2016.

To illustrate, consider the example of *Janakpuri*, which had a station prior to 2016, while *Dabri* and *Bindapur* did not, compelling residents to travel to *Janakpuri*. The first station

[26]There is one police jurisdiction which was treated in the fourth quarter of 2019. However, due to lack of post-treatment data for this jurisdiction, I include it in the never-treated group for the analysis.

Table 1: Quarterly Number of Treated Jurisdictions

Quarters	2017 Q2	2017 Q4	2018 Q1	2018 Q2	2018 Q3	2018 Q4
Treated Jurisdictions	2	3	4	7	2	6

Notes: This table presents the quarterly number of jurisdictions treated between 2016 and 2019.

opened in *Dabri* during the analysis period. Consequently, the jurisdiction of *Dabri* Police Station is treated, while that of *Bindapur* Police Station falls into the control group and that of *Janakpuri* Police Station into the already treated group. Figure A8 in Appendix Section A.3 illustrates this example.

The recent developments in the econometrics of difference-in-differences (DiD) for staggered treatment setting reveal that the dynamic two way fixed effects (TWFE) regression can yield biased estimates of the actual average treatment effect on the treated (ATT) due to heterogeneous treatment effects across different opening dates and police jurisdictions. This stems from the estimate including 'forbidden' comparisons of later treated jurisdictions with the earlier treated ones as the control group (Sun & Abraham, 2021).[27] To assess the extent of this issue in my context, I employ the diagnostic test proposed by Goodman-Bacon (2021), revealing that 4% of the TWFE estimate originates from these 'forbidden' comparisons. The results are reported in Table 2. The literature also recommends exploring alternative estimators even if TWFE does not pose a problem, given that the control group for TWFE estimation lacks transparency and the weights used for averaging ATTs might not be policy-relevant.

To address these concerns, I adopt an alternative estimation and inference approach discussed in the recent econometrics literature. Specifically, I utilize the group-time-specific average treatment effects (ATTs) proposed in Callaway & Sant'Anna (2021), which eliminates these 'forbidden' comparisons. In particular, the ATT for each treated cohort g in each quarter t is identified as follows:

[27]For a synthesis of the recent DiD literature, see Roth et al. (2023).

Table 2: Goodman-Bacon Decomposition of Static TWFE Estimates

Comparison	Weight	Estimate
Treated vs Never-Treated	0.903	−2.546
Earlier vs Later Treated	0.058	−1.031
Later vs Earlier Treated	0.039	0.207
N	1,264	1,264

Notes: This table presents the decomposition of the static TWFE estimates for reported HoS per 100,000 population following Goodman-Bacon (2021). The comparison group includes never-treated jurisdictions.

$$\mathrm{CAW}_{it} = \alpha_t^{g,t} + \alpha_i^{g,t} + \beta^{g,t}(\mathrm{G}_g \times \mathbb{1}\{T = t\}) + \epsilon_{it} \tag{1}$$

where CAW_{it} is the reported crimes per 100,000 people at police station i in quarter t.[28] T is the set of all the quarters in the panel. G_g is a dummy variable which takes the value 1 if police jurisdiction i was treated in quarter g, else 0. α_i and α_t capture the police jurisdiction-level fixed effects and time-fixed effects, respectively, accounting for unobserved determinants of reported crime and time trends in such crimes. Standard errors are clustered at the police jurisdiction level. The coefficient of interest is $\beta^{g,t}$, which measures the ATT for each police station i in each treatment cohort g in each quarter t.

The dynamic treatment effect is computed by aggregating these $\mathrm{ATT}_{g,t}$ by the length of exposure e of a police jurisdiction i to the treatment as follows:

$$\theta(e) = \Sigma_{g \in \mathrm{G}_g} \mathbb{1}\{g + e <= T\}) P(\mathrm{G}_g = g | g + e <= T) \mathrm{ATT}(g, g + e) \tag{2}$$

This facilitates the analysis of the treatment effect dynamics across different aggregations - whether average treatment effects of the opening of a metro station in a police jurisdiction vary with the length of exposure e to the treatment or across treatment cohorts G_g and whether the cumulative average treatment effect across all the police stations until some time $\tilde{t}$ varies with $\tilde{t}$. In the Sub-section 5.2, I elucidate the rationale for using the esti-

[28]These are calculated using the FIR data.

mation strategy provided by Callaway & Sant'Anna (2021) over other alternatives (Sun & Abraham (2021), de Chaisemartin & D'Haultfœuille (2020) and Borusyak et al. (2021)) and demonstrate the robustness of my results when using these other alternatives.

Callaway & Sant'Anna (2021) allow for the estimation to incorporate two types of control groups. The first is the 'not-yet-treated' group, encompassing all those police jurisdictions that have yet to receive a metro station within the analysis period and those that have not received one in the analysis period. The second is the 'never treated' group, comprising 55 jurisdictions that did not receive a metro station during the study period but may or may not get one post the analysis period. In my baseline specification, I use the 'not-yet-treated' group as my control group and a period of six quarters before and after the treatment. To fortify the integrity of my findings, I validate the robustness of my outcomes by comparing results with an alternate control group consisting only of the 55 'never treated' jurisdictions.

Finally, despite the time lag between the decision to construct a metro station and its actual opening, I assume that a jurisdiction gets treated only after the station's opening. This assumption ensures that causal effects are absent prior to station commencement. It is important to note that Delhi's metro system, apart from improving connectivity across different areas of Delhi, was designed to address escalating challenges posed by a rising number of registered motor vehicles and population growth on local infrastructure.[29] The stations were planned and built based on the geographic, administrative and logistical factors, rather than area-specific crime statistics.[30] This suggests that endogeneity might not be a concern here. So, the main identifying assumption is that in the absence of the inaugural metro station in the treated jurisdictions, the crime outcomes for a group treated in quarter G_g and the groups remaining without a metro station until quarter G_g will exhibit parallel trends (Callaway & Sant'Anna, 2021). In the subsequent section, I examine the validity

[29] https://www.centreforpublicimpact.org/case-study/construction-delhi-metro; https://housing.com/news/delhi-metro-phase-iv-finally-approved-government/

[30] Geographic factors include the feasibility of building an underground network - https://www.ice.org.uk/what-is-civil-engineering/what-do-civil-engineers-do/delhi-metro/. Administrative factors include the speed of getting necessary approvals, and logistical factors include the land investigations of each construction site - https://themetrorailguy.com/delhi-metro-phase-4-information-map/.

of the parallel trends using the dynamic treatment effect coefficients estimated through the strategy outlined in Callaway & Sant'Anna (2021).

5 Results

5.1 Main Results

The outcome variables are the different categories of reported per capita crimes per 100,000 population. Table 3 represents the summary statistics for the crime categories for the 24 treated jurisdictions and 55 never-treated jurisdictions. Table 4 presents the weighted average of dynamic group-time specific ATTs, standard errors clustered at the police jurisdiction level, and the associated p-values, using the approach outlined in Callaway & Sant'Anna (2021).[31] Figure 2 illustrates the corresponding event study plots. The coefficients in the period prior to the opening of the first metro station are statistically insignificant, lending support to the plausibility of the parallel trends between the treated and the control jurisdictions in the absence of the metro station.[32]

Table 3: Summary Statistics for Reported Crimes for All Jurisdictions (2016-2019)

Reported Crimes per 100k Population	N	Mean	Std. Dev.	Minimum	Maximum
HoS crimes against women	1,264	6.635	6.119	0	82.719
Non-gender-specific street crimes	1,264	49.511	47.033	0	457.020
All crimes against women	1,264	12.892	8.415	0	82.719
Domestic violence	1,264	4.838	4.417	0	41.359

Notes: This table presents the summary statistics for the different reported crime categories that form the outcome variable. These variables are calculated per 100,000 population at the police-station quarter level.

First, I analyse the impact of opening the first metro station in a police jurisdiction on reported HoS crimes perpetrated on women. Column 1 of Table 4 reports that relative

[31] I use the Stata module, *csdid* (Rios-Avila et al., 2023), to implement Callaway & Sant'Anna (2021).

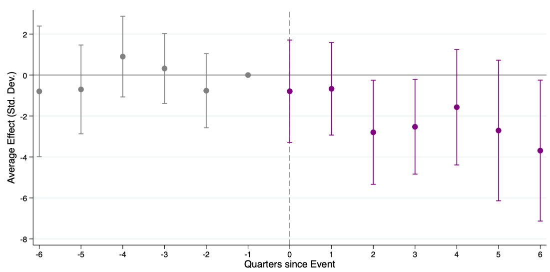
[32] I intend to correct for potential bias from differences in pre-treatment trends following Roth (2022). I also plan to assess the sensitivity of my results to possible violations of parallel trends assumption following Rambachan & Roth (2023).

Table 4: Main Results: Comparison Group - Never and Not-yet-treated Jurisdictions

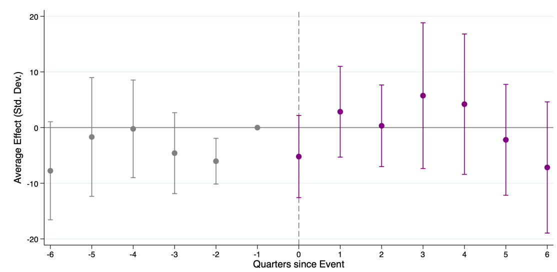
	HoS Crimes Against Women	Non-Gender-Specific Street Crimes	All Crimes Against Women	Domestic Violence
Post 1^{st} Metro Station	$\mathbf{-1.947}^{*}$	4.426	−2.140	0.123
	(1.099)	(5.078)	(1.372)	(0.632)
Observations	1,264	1,264	1,264	1,264
Jurisdiction Fixed Effects	Yes	Yes	Yes	Yes
Quarter Fixed Effects	Yes	Yes	Yes	Yes
P-value	0.077	0.383	0.119	0.846

Notes: This table presents the dynamic ATT following Callaway & Sant'Anna (2021) of the opening of the first metro station in a police jurisdiction. The comparison group includes not-yet-treated and never-treated jurisdictions. The time frame consists of six quarters before and after the treatment. Standard errors are clustered at the police jurisdiction level. Fixed effects are for each police jurisdiction and each quarter between 2016 and 2019. ***, **, and * indicate significance at the 1, 5, and 10 percent critical levels.

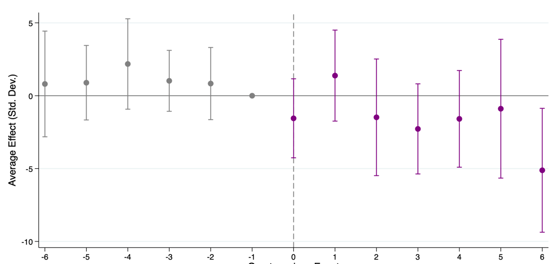
Figure 2: Even Study Plots: Comparison Group - Never and Not-yet-treated Jurisdictions



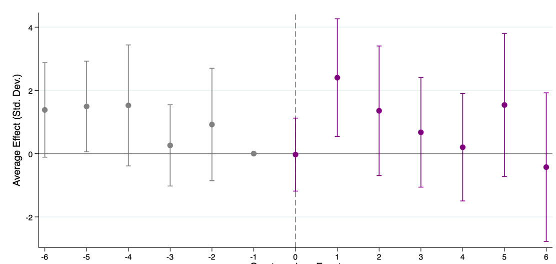
(a) **HoS Crimes Against Women**



(b) Non-Gender-Specific Street Crimes



(c) All Crimes Against Women



(d) Domestic Violence

Notes: This figure presents the event study coefficients and the 95% confidence interval following Callaway & Sant'Anna (2021) of the opening of the first metro station in a police jurisdiction. The comparison group includes not-yet-treated and never-treated jurisdictions. The time frame consists of six quarters before and after the treatment. Standard errors are clustered at the police jurisdiction level. Fixed effects are for each police jurisdiction and each quarter between 2016 and 2019.

to the 'not-yet-treated' police jurisdictions, reported HoS crimes against women exhibit a reduction of approximately 2 reported HoS crimes per 100,000 population following the opening of the first metro station. This quantifies to a 29.34% reduction in reported HoS crimes against women in the treated jurisdictions, compared to the control group, and is statistically significant at a p-value of 0.077.

However, it might be the case that this decrease in reported HoS crimes against women is simply a part of an overall decrease in reported crimes in public places following the opening of a metro station. Such a decline might stem from an overall enhancement in safety around the metro station area. So, I analyse the impact of first metro stations on non-gender-specific street crimes.[33] The result, in Column 2 of Table 4, indicates no statistically significant relationship between these crimes and the opening of the first metro station within the jurisdiction. Columns 3 and 4 of Table 4 show that compared to the control group, reported cases of crimes against women and domestic violence show no significant change in the quarters after the opening of the metro station, suggesting it is the sexual harassment crimes perpetrated on women in public places which show a reduction.[34]

Furthermore, it is also possible that this decrease in reported HoS crimes against women might be a result of these crime reports shifting to incidents within the metro rail or station premises. As discussed in Sub-section 3.2, crimes occurring in the metro rail or the station are reported at the metro police station, under whose purview the particular metro station falls. This means that the reduction in these cases at general police stations might be accompanied by an increase in similar cases at metro police stations. To explore this possibility, I examine the crime reports from those metro police stations where a new metro station was integrated into their jurisdiction between 2016 and 2019. During this entire period, a total of 42 HoS

[33]Non-gender specific street crimes include all the possible crimes that can occur in the public spaces against any gender and exclude all the crimes like sexual harassment, which are perpetrated mostly on female gender. Table A1 in Appendix Section A.1 lists the sections from the Indian Penal Code 1860 used to generate this crime category.

[34]Approximately 90% of reported crimes against women consist of domestic violence incidents or those occurring in public spaces. The rest 10% are kidnapping and rape cases. Table A1 in Appendix Section A.1 discusses the sections from the Indian Penal Code 1860 used to create these categories.

crimes against women were reported at these metro police stations. However, the total treatment effect of the opening of metro stations is a reduction of 271 reported HoS crimes against women at general police stations. This initial analysis suggests that displacement of HoS crimes might not be occurring.[35]

5.2 Robustness

I conduct a series of thorough robustness checks to enhance the credibility of my findings. These include employing an alternative comparison group and applying different estimation methods from the recent DiD literature.

5.2.1 Alternate Comparison Group

As discussed in Section 4, I reinforce the resilience of my results by contrasting the outcomes obtained from an alternate control group comprising 55 'never-treated' jurisdictions. Table A3 in Appendix Section A.4 shows that the results remain robust to this reference group. Specifically, the reported incidents of HoS crimes against women decrease by approximately 2 reported cases per 100,000 population, a decrease of 29.59% when compared to the control group, and is statistically significant at a p-value of 0.073.

5.2.2 Other Estimation Methods

I have adopted the approach in Callaway & Sant'Anna (2021) over other estimation methods found in recent econometrics literature focusing on even study estimation. This decision is grounded in several key considerations, which I outline below.

To begin with, the primary distinction between Callaway & Sant'Anna (2021) and Sun & Abraham (2021) lies in the comparison group. Unlike Callaway & Sant'Anna (2021),

[35] Currently, these statistics are presented at the level of total reports rather than per 100,000 population. Additionally, as all metro police stations had at least one metro station under their jurisdiction before 2016, establishing a distinct pre- and post-period requires further textual analysis. Consequently, this conclusion should be approached with caution. I intend to provide more rigorous evidence once I compile a comprehensive data set of crime reports, alongside metro ridership data broken down by different stations, quarters, and gender demographics.

Sun & Abraham (2021) does not facilitate a comparison with the 'not-yet-treated' group. Instead, it offers only two alternative control groups, the 'last treated' or the 'never treated' groups. Next, the strategy in Borusyak et al. (2021) introduces a limitation that affects its applicability in my study. Specifically, this approach yields unreliable estimates and introduces a downward bias in the standard deviation when the count of treated observations for a particular temporal point about the treatment quarter falls below 30. Furthermore, the estimation strategy in de Chaisemartin & D'Haultfœuille (2020) is not able to estimate pre-treatment effects for more extended periods (de Chaisemartin & D'Haultfoeuille, 2023). Consequently, given my specific circumstances, I do not employ these estimation strategies.

An additional merit of employing the methodology proposed by Callaway & Sant'Anna (2021), as elaborated upon in Section 4, is its ability to facilitate the examination of treatment effect dynamics across diverse aggregations. This encompasses investigating whether the average treatment effects resulting from the opening of a metro station within a given police jurisdiction exhibit variations based on the length of exposure to the treatment. It also enables the assessment of potential disparities across different treatment cohorts. Furthermore, it supports exploring how the cumulative average treatment effect across all police jurisdictions up until a specific point in time may display fluctuations with time.

In light of these considerations, I have rigorously examined the robustness of my findings using these latest estimation strategies and the dynamic Two-Way Fixed Effects (TWFE) strategy, with the 'never treated' group serving as the control.[36] The results presented in Table A4 and Figure A11 in the Appendix Section A.4 show that the magnitude of the estimates and their standard errors remain stable across these estimation approaches.[37]

[36] I use the Stata modules - *eventstudyinteract* (Sun, 2021), *did_imputation* (Borusyak, 2021), *did_multiplegt* (de Chaisemartin et al., 2019) and *reghfde* (Correia, 2014) - to implement the estimation strategies proposed in Sun & Abraham (2021), Borusyak et al. (2021) and de Chaisemartin & D'Haultfœuille (2020), and dynamic TWFE strategy, respectively.

[37] I do not report estimates from Borusyak et al. (2021) because the estimates may be unreliable and the standard errors may be downward biased due to insufficient sample size.

6 Mechanisms

6.1 Safety Effect

Evidence using reported HoS crimes against women on the public bus network: Well-illuminated metro stations and rail systems, equipped with security personnel and surveillance cameras ensuring safety during operational hours, may present a more secure public transport alternative compared to existing options. This heightened security presence has the potential to act as a crime deterrent, thereby reducing reported incidents of crimes against women during their journey.

As discussed in Section 3.2, Delhi Metro categorizes stations by different lines. Between 2016 and 2019, new stations were inaugurated along the magenta, pink, and violet lines.[38] The newly opened magenta line stations facilitated improved connectivity between the South-West and South-East regions of Delhi, reducing commuting times by nearly half. Figure 3 shows an illustration of this travel time reduction for two neighbourhoods in South-West and South-East Delhi. The consequential impact of this development was evident in the subsequent surge in metro ridership, as illustrated in Figure 4.

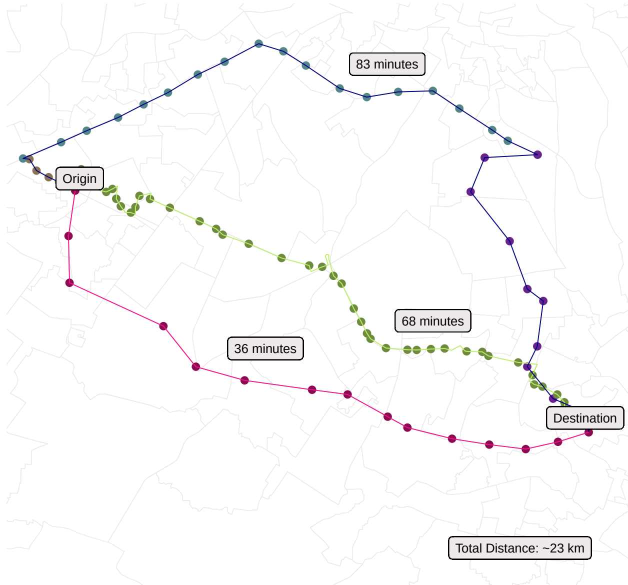
I further divided the treatment into three groups based on the new stations along these three lines and estimate the following TWFE regression equation:

$$\text{HoS}_{it} = \beta_1 \text{Magenta}_i \text{D}_{it} + \beta_2 \text{Pink}_i \text{D}_{it} + \beta_3 \text{Violet}_i \text{D}_{it} + \alpha_i + \alpha_t + \epsilon_{it} \tag{3}$$

where HoS_{it} is the reported HoS crimes against women per 100,000 population at police station i in quarter t. Magenta_i, Pink_i and Violet_i are dummy variables that take the value 1 for all the jurisdictions where the first metro station that opened was on magenta, pink and violet line, respectively and 0 otherwise. D_{it} is a dummy variable that takes the value 1 for

[38]Figure A9 in Appendix Section A.3 illustrated the stations that opened on these three lines. https://timesofindia.indiatimes.com/city/noida/noida-to-south-delhi-by-metro-in-16-minutes-when-magenta-line-opens/articleshow/61124513.cms#:~:text=Noida%3A%20The%20Magenta%20Line%20will,from%20Kalkaji%20Mandir%2DBotanical%20Garden.

Figure 3: Travel Time Reduction between South-West and South-East Delhi - An Illustration



Notes: This figure presents alternative mass public transport options between two neighbourhoods in South-West (Origin) and South-East (Destination) Delhi, covering 23 kilometres of distance. The blue and violet lines show the route traversed through the stations on these lines taking 83 minutes, the green line shows the bus route taking 68 minutes, and the magenta line shows the minimum travel time of 36 minutes, which became possible after the opening of stations on the magenta line.

Figure 4: Quarterly Metro Ridership

Notes: This figure presents quarterly metro ridership. The violet, magenta and pink lines depict the opening quarters of stations on these lines, respectively. *Source:* `https://data.gov.in/search?title=delhi%20metro`

Table 5: Heterogeneous Treatment Effect by Metro Lines

	HoS Crimes Against Women
Magenta Line × Post	-3.297^{**}
	(1.653)
Pink Line × Post	-1.920^{*}
	(1.032)
Violet Line × Post	-1.120
	(2.486)
Observations	1,264
Jurisdiction Fixed Effects	Yes
Quarter Fixed Effects	Yes

Notes: This table presents the heterogeneity in the static TWFE estimates of the opening of the first metro station in a police jurisdiction based on magenta, pink and violet metro lines. The comparison group includes not-yet-treated and never-treated jurisdictions. Standard errors are clustered at the police jurisdiction level. Fixed effects are for each police jurisdiction and each quarter between 2016 and 2019. ***, **, and * indicate significance at the 1, 5, and 10 percent critical levels.

all the quarters after the treatment in quarter G_g. α_i and α_t capture the police jurisdiction-level fixed effects and time-fixed effects, respectively. Standard errors are clustered at the police jurisdiction level. The results presented in Table 5 show that the effect in column 3 of Table 4 is most pronounced for jurisdictions where the first metro station that opened was on the magenta line.

The rise in metro ridership following the opening of stations on the magenta line, coupled with a concurrent decrease in reported harassment and assault crimes against women, offers suggestive evidence in favour of the metro being a safer mode of public transport. To further assess the metro's safety in comparison to other modes, I focus on crimes occurring on buses, the primary alternative mode of mass public transport. Textual analysis of crime reports enables the identification of incidents on buses or at bus stops, including details such as the specific route, bus number and name of the bus stop.[39] Table 6 presents the results from estimating a TWFE regression equation analogous to Equation 3, with reported HoS crimes against women in buses or at bus stops as the dependent variable. For all the

[39]Figure A6 in Appendix Section A.1 illustrates a snapshot of these details in the crime report registered for a crime occurring on a bus.

Table 6: Reported HoS Crimes Against Women on the Bus Network

	HoS Crimes Against Women
Magenta Line × Post	-0.943^{***}
	(0.307)
Pink Line × Post	-0.634^{*}
	(0.362)
Observations	918
Jurisdiction Fixed Effects	Yes
Quarter Fixed Effects	Yes

Notes: This table presents the heterogeneity in the static TWFE estimates of the opening of the first metro station in a police jurisdiction based on magenta, pink and violet metro lines on reported HoS crimes perpetrated on women in buses or at bus stops. The comparison group includes not-yet-treated and never-treated jurisdictions. Standard errors are clustered at the police jurisdiction level. Fixed effects are for each police jurisdiction and each quarter between 2016 and 2019. ***, **, and * indicate significance at the 1, 5, and 10 percent critical levels.

police jurisdictions where the first metro station opened along the magenta and the pink line, reported HoS crimes against women on the alternative mode of mass public transport reduced significantly. This reduction, combined with the decrease in modal split of buses and the increase in that of metros between 2007 and 2018, provides suggestive evidence of substitutability between the two modes of public transport.[40] These findings provide initial suggestive evidence supporting the *safety* effect of the metro rail system.

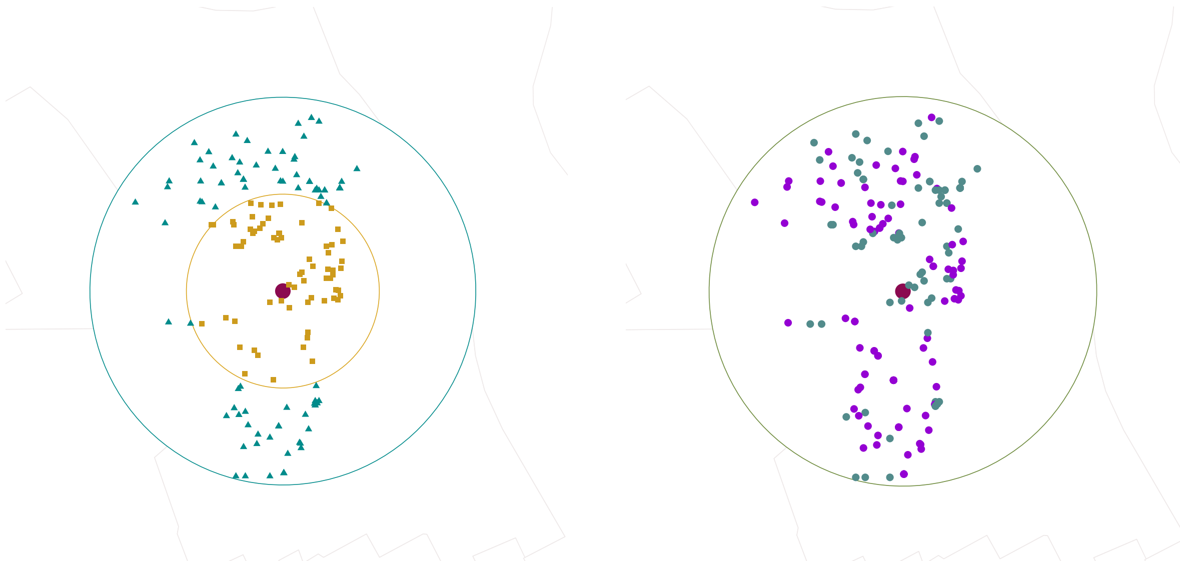
Evidence using SafetiPin data: I use area safety data from SafetiPin, a mobile safety app in Delhi (Viswanath & Basu, 2015).[41] The SafetiPin app allows users to conduct safety audits of their current location, rating it based on nine parameters, namely, the extent of lighting, openness of the space, visibility, presence of people (especially women and children), presence of security personnel and surveillance cameras, condition of the walking path, availability of public transport, and the general feel of the area. Based on these parameters, a composite safety score is assigned to each location on a scale from 0 (least safe) to 10 (safest). My dataset includes 15,131 audits conducted between 2013 and 2024.[42]

[40] Figure A12 in Appendix Section A.5 shows this modal split.

[41] `https://safetipin.com`

[42] The data was shared by the SafetiPin team.

Figure 5: SafetiPin Audits - An Illustration



(a) Treatment and Control (b) Pre- and Post-Metro Station in a 1km Radius

Notes: Figure (a) presents an illustration of all the audit points within a 500 metre radius of a metro station (treatment) and all the audit points lying between 500 and 1,000 metres of that metro station (control). Figure (b) depicts all the points within a 1,000 metres radius of a metro station, where purple points correspond to all the audits conducted before the opening of that metro station, and green points correspond to those after the opening of the metro station.

I focus on all metro stations that opened between 2013 and 2019. For each station, I draw two circles with a radius of 500 meters and 1 kilometer around the station. Audit points within the 500 meter radius are classified as the treatment group, while those located between 500 meters and 1 kilometer form the control group.[43] Figure 5 illustrates one such example. Using the station opening date as the event, I run a TWFE regression. As shown in Table 7, the safety score of areas within 500 meters of a metro station increases by an average of 0.59 points compared to areas within 500 meters to 1 kilometer, relative to a baseline average of 7.22 points, after the station opens (column 1). Additionally, when considering all audit points within the 1 kilometer radius, the safety score decreases by an average of 1.42 points as the distance to the nearest metro station increases (column 2). This provides further suggestive evidence supporting the *safety* effect of metro stations.

[43] On average, people in Delhi walk 500 meters to access a metro station (Advani & Tiwari, 2005). The average distance between two metro stations is 2 kilometers.

Table 7: Evidence using SafetiPin Data

	Safety Score	
	0−500−1000m	≤1000m
Treatment × Post	0.589***	
	(0.193)	
Distance × Post		−1.416***
		(0.407)
Treatment	0.220*	
	(0.126)	
Distance		−0.243
		(0.260)
Observations	1,264	1,264
Ward Fixed Effects	Yes	Yes
P-value	0.003	0.001

Notes: This table presents the TWFE estimates of the opening of a metro station. The treatment group is all the audit points within 500 metres of the new metro station. The control group is all the audit points within 500 and 1,000 metres of the new metro station. Robust standard errors are used. Ward-level fixed effects are added. ***, **, and * indicate significance at the 1, 5, and 10 percent critical levels.

Evidence using survey data: Between October and November 2012, the International Centre for Research on Women (ICRW) conducted a survey for the Safe Cities initiative in Delhi. This initiative was a collaboration between the Delhi Government, UN Women, and the NGO Jagori. ICRW surveyed a sample of 2,001 females and 1,003 males aged between 16 and 49 years from seven census tracts. The survey revealed that 55%, 57%, 51%, and 46% of respondents perceived bus stops, streets, markets, and parks as unsafe, respectively. In contrast, only 20% of respondents reported metro stations to be unsafe.[44]

The Observer Research Foundation, an India-based think tank, conducted a survey on the safety of travel modes between December 2019 and September 2020. The survey included 3,537 female respondents from the 15 most populous cities in India. Only 12% of respondents reported buses as safe, while 48% identified the metro or train as the safest mode of transport.[45] These survey findings further reinforce the evidence supporting the *safety* effect

[44] `http://www.safedelhi.in/sites/default/files/reports/Safety-of-Women-in-Delhi.pdf`.

[45] `https://www.orfonline.org/wp-content/uploads/2021/05/ORF_Monograph_WomenOnTheMove.pdf`.

of the metro system.

6.2 Volume Effect

As discussed in Section 2, improved transportation access might lead to an increase in mobility due to improved connections to different parts of the city. As a result, the victim pool in the treated jurisdictions might increase, leading to an increase in criminal opportunities in the jurisdictions with the new metro station. However, the rise in metro ridership, as illustrated in Figure 4, and the decrease in reported harassment and assault crimes against women, as indicated in column 1 of Table 4, seem to suggest that the *volume* effect may be getting overpowered by the *safety* effect. To delve deeper into this dynamic, my next step involves acquiring ridership data for buses. Additionally, I plan to segregate the metro and bus ridership data by gender, conducting a thorough analysis of ridership trends. This approach will help me furnish additional evidence and insights into the extent to which the *volume* effect contributes to these observed dynamics.

6.3 Reporting Effect

The primary challenge when analyzing crime data lies in accurately measuring the occurrences of actual criminal incidents. As discussed in Section 2, improved access to education and employment opportunities empowers women and makes them financially independent, potentially making them less tolerant of any form of transgressions and more inclined to report such incidents. Additionally, the opening of metro stations might result in heightened police vigilance and increased arrests, thereby making reporting a more effective option Bhuller et al. (2013). However, the observed trend suggests a decline in reported HoS crimes against women following the opening of metro stations, while domestic violence remains unchanged. Moreover, non-gender-specific street crimes display stability. As a result, the *reporting* effect does not seem to be at work here.

To bolster the validity of my findings, I will utilise the dates of report filing and the occurrence of crimes, along with reasons for potential reporting delays, to apply the duration model, as proposed by Gauthier (2022). This analysis will help differentiate between the influence of actual crime incidence and reported crimes. If the number of crimes reported with a delay remains consistent in the post-treatment quarters, it suggests that the observed effect is likely due to actual crime occurrences. Conversely, an increase in such delayed reports might indicate a *reporting* effect being at play.

6.4 Spatial Effect

The impact of a new metro station within a specific police jurisdiction often ripples beyond its immediate borders, affecting neighbouring areas. The convenience of metro services might attract females from the neighbouring jurisdictions to the treated jurisdictions, drawn by the convenience of metro services, leading to a surge in overall criminal incidents within the treated jurisdictions. As a result, the negative impact on the reported HoS crimes against women in Table 4 is less negative.

To assess the presence of this spatial effect on the treated jurisdictions, I will geocode the crime locations to census boundaries called 'wards'. Then, I will remove the wards that share a border with the treated wards to create a new control group (Kline & Moretti, 2013). I will then compare the treated wards with this control group to analyse whether the initially observed negative impact on the reported HoS crimes against women in Table 4 becomes more pronounced.

7 Conclusion

In addition to its role in promoting social inclusion and facilitating connectivity, the question arises whether investment in mass transit impacts crimes against women, which is a severe worldwide phenomenon. In my investigation, I delve into the expansion of the metro rail system in Delhi, India, which is notorious for being unsafe for women and where travel connectivity is sparse, creating an obstacle to women's economic development.[46] To achieve this, I curate an original data set of crime reports at the granularity of individual police stations using *Optical Character Recognition* and *Named Entity Recognition*.

By leveraging the exogenous variation in the staggered opening of the first metro stations across different police jurisdictions in Delhi between 2016 and 2019, my analysis reveals a 29 percent reduction in reported incidents of sexual harassment against women in public spaces. However, while this decrease pertains to reported sexual harassment, reported cases of domestic violence and overall crimes against women demonstrate no significant shifts. Importantly, the observed decline in reported sexual harassment is not manifested in an overall reduction in reported non-gender-specific street crimes or substituted by an increase in reported sexual harassment within metro stations or rails.

These findings uncover a nuanced relationship between the expansion of metro infrastructure and specific forms of gender-based crimes. While decreased reported sexual harassment is a significant outcome, the broader landscape of women-related crimes remains relatively unaffected. This insight prompts a deeper exploration into the mechanisms underlying the observed dynamics, namely, safety, volume, spatial, and reporting effects, emphasizing the intricate nature of the issue.

In addition to thoroughly disentangling these effects, I am also geocoding crime locations to gain a more nuanced understanding of the impact of new metro stations on reported crimes against women.

[46] https://edition.cnn.com/travel/article/worst-transport-for-women/index.html; https://timesofindia.indiatimes.com/city/delhi/delhi-government-took-slew-of-steps-to-ensure-womens-safety/articleshow/74732463.cms

References

Advani, M., & Tiwari, G. (2005). Evaluation of public transport systems: Case study of delhi metro.. Retrieved from https://api.semanticscholar.org/CorpusID:18253744

Amaral, S., Bhalotra, S., & Prakash, N. (2021). Gender, crime and punishment: Evidence from women police stations in india. Retrieved from https://ssrn.com/abstract=3827615

Bhuller, M., Havnes, T., Leuven, E., & Mogstad, M. (2013, 04). Broadband Internet: An Information Superhighway to Sex Crime? The Review of Economic Studies, 80(4), 1237-1266. Retrieved from https://doi.org/10.1093/restud/rdt013

Billings, S. B., Leland, S., & Swindell, D. (2011). The effects of the announcement and opening of light rail transit stations on neighborhood crime. Journal of Urban Affairs, 33(5), 549-566. Retrieved from https://doi.org/10.1111/j.1467-9906.2011.00564.x

Borker, G. (2021). Safety first : Perceived risk of street harassment and educational choices of women. Policy Research Working Paper, 9731(1). Retrieved from https://openknowledge.worldbank.org/handle/10986/36004

Borusyak, K. (2021, June). Did_imputation: Stata module to perform treatment effect estimation and pre-trend testing in event studies. Statistical Software Components, Boston College Department of Economics. Retrieved from https://ideas.repec.org/c/boc/bocode/s458957.html

Borusyak, K., Jaravel, X., & Spiess, J. (2021, August). Revisiting Event Study Designs: Robust and Efficient Estimation (Papers). Retrieved from https://ideas.repec.org/p/arx/papers/2108.12419.html

Callaway, B., & Sant'Anna, P. H. (2021). Difference-in-differences with multiple time periods. Journal of Econometrics, 225(2), 200-230. Retrieved from https://www.sciencedirect.com/science/article/pii/S0304407620303948

Card, D., & Dahl, G. B. (2011, 02). Family Violence and Football: The Effect of Unexpected Emotional Cues on Violent Behavior*. The Quarterly Journal of Economics, 126(1), 103-143. Retrieved from https://doi.org/10.1093/qje/qjr001

Correia, S. (2014, July). Reghdfe: Stata module to perform linear or instrumental-variable regression absorbing any number of high-dimensional fixed effects. Statistical Software Components, Boston College Department of Economics. Retrieved from https://ideas.repec.org/c/boc/bocode/s457874.html

de Chaisemartin, C., & D'Haultfoeuille, X. (2023). Difference-in-differences estimators of intertemporal treatment effects.

de Chaisemartin, C., d'Haultfoeuille, X., & Guyonvarch, Y. (2019). Did_multiplegt: Stata module to estimate sharp difference-in-difference designs with multiple groups and periods (SciencePo Working papers Main No. hal-03946768). HAL. Retrieved from https://ideas.repec.org/p/hal/spmain/hal-03946768.html

de Chaisemartin, C., & D'Haultfœuille, X. (2020, September). Two-way fixed effects estimators with heterogeneous treatment effects. American Economic Review, 110(9), 2964-96. Retrieved from https://www.aeaweb.org/articles?id=10.1257/aer.20181169

Dominguez Gonzalez, K., Machado, A. L., Bianchi Alves, B., Raffo, V., Guerrero, S., & Portabales, I. (2020). Why does she move? World Bank. Retrieved from https://elibrary.worldbank.org/doi/abs/10.1596/33466

Gauthier, G. (2022, 10). Measuring crime reporting and incidence: Method and application to #metoo. Retrieved from https://ssrn.com/abstract=4242506

Goodman-Bacon, A. (2021). Difference-in-differences with variation in treatment timing. Journal of Econometrics, 225(2), 254-277. Retrieved from https://www.sciencedirect.com/science/article/pii/S0304407621001445

Iyer, L., Mani, A., Mishra, P., & Topalova, P. (2012, July). The power of political voice: Women's political representation and crime in india. American Economic Journal: Applied Economics, 4(4), 165-93. Retrieved from https://www.aeaweb.org/articles?id=10.1257/app.4.4.165

Jassal, N. (2020). Gender, law enforcement, and access to justice: Evidence from all-women police stations in india. American Political Science Review, 114(4), 1035–1054.

Khanna, G., Medina, C., Nyshadham, A., Ramos-Menchelli, D., Tamayo, J., & Tiew, A. (Working Paper). Spatial mobility, economic opportunity, and crime.

Kline, P., & Moretti, E. (2013, 11). Local Economic Development, Agglomeration Economies, and the Big Push: 100 Years of Evidence from the Tennessee Valley Authority *. The Quarterly Journal of Economics, 129(1), 275-331. Retrieved from https://doi.org/10.1093/qje/qjt034

Luca, D. L., Owens, E., & Sharma, G. (2015). Can alcohol prohibition reduce violence against women? American Economic Review, 105(5), 625-29. Retrieved from https://EconPapers.repec.org/RePEc:aea:aecrev:v:105:y:2015:i:5:p:625-29

McDougal, L., Krumholz, S., Bhan, N., Bharadwaj, P., & Raj, A. (2021). Releasing the tide: How has a shock to the acceptability of gender-based sexual violence affected rape reporting to police in india? Journal of Interpersonal Violence, 36(11-12), NP5921-NP5943. Retrieved from https://doi.org/10.1177/0886260518811421

Miller, A. R., & Segal, C. (2018, 09). Do Female Officers Improve Law Enforcement Quality? Effects on Crime Reporting and Domestic Violence. The Review of Economic Studies, 86(5), 2220-2247. Retrieved from https://doi.org/10.1093/restud/rdy051

Rambachan, A., & Roth, J. (2023, 02). A More Credible Approach to Parallel Trends. The Review of Economic Studies. Retrieved from https://doi.org/10.1093/restud/rdad018

Rios-Avila, F., Sant'Anna, P., & Callaway, B. (2023). Csdid: Stata module for the estimation of difference-in-difference models with multiple time periods. Retrieved from https://EconPapers.repec.org/RePEc:boc:bocode:s458976

Roth, J. (2022, September). Pretest with caution: Event-study estimates after testing for parallel trends. American Economic Review: Insights, 4(3), 305-22. Retrieved from https://www.aeaweb.org/articles?id=10.1257/aeri.20210236

Roth, J., Sant'Anna, P. H. C., Bilinski, A., & Poe, J. (2023). What's trending in difference-in-differences? a synthesis of the recent econometrics literature.

Siddique, Z. (2022). Media reported violence and female labor supply. Economic Development and Cultural Change, 70(4), 1337–1365.

Sun, L. (2021, August). Eventstudyinteract: Stata module to implement the interaction weighted estimator for an event study. Statistical Software Components, Boston College Department of Economics. Retrieved from https://ideas.repec.org/c/boc/bocode/s458978.html

Sun, L., & Abraham, S. (2021). Estimating dynamic treatment effects in event studies with heterogeneous treatment effects. Journal of Econometrics, 225(2), 175-199. Retrieved from https://www.sciencedirect.com/science/article/pii/S030440762030378X

Viswanath, K., & Basu, A. (2015). Safetipin: an innovative mobile app to collect data on women's safety in indian cities. Gender and Development, 23(1), 45–60. Retrieved 2024-08-27, from http://www.jstor.org/stable/24697524

Viswanath, K., & Mehrotra, S. T. (2007). 'shall we go out?' women's safety in public spaces in delhi. Economic and Political Weekly, 42(17), 1542–1548. Retrieved 2022-11-12, from http://www.jstor.org/stable/4419521

A Appendix

A.1 Reported Crimes against Women

In this section, I will elaborate on my meticulous data preparation strategy to curate an exhaustive data set of First Information Reports (FIRs) in Delhi. The primary goal is to transform raw and unstructured data into a machine-readable format for comprehensive analysis.

The initial step entailed web scraping all available FIRs from the official website of the Delhi Police, a process executed through the command line utility tool *wget*.[47] A sample FIR extracted through web scraping is illustrated in Figure A1.

[47] `http://59.180.234.21:8080/citizen/firSearch.htm`

Figure A1: A Sample FIR

FIRST INFORMATION REPORT
(Under Section 154 Cr.P.C.)

(धारा 154 दंड प्रक्रिया सहिंता के तहत)

1. District (ज़िला):SOUTH **P.S.(थाना):** MEHRAULI **Year(वर्ष):** 2016 **FIR No(प्र.सू.रि.सं.):**0136 **Date :** 15/01/2016

2. Act(s)(अधिनियम): **Section(s)(धारा(एँ)):**

- IPC 1860 451/506/509/34

3. Occurrence of Offence (अपराध की घटना):
(a) Day(दिन): **Date From(दिनांक से):** **Date To(दिनांक तक):**
Time Period (समय अवधि): **Time From (समय से):** **Time To (समय तक):**
(b)Information received at P.S.(थाना जहां सूचना प्राप्त हुई): **Date(दिनांक):** 15/01/2016 **Time (समय):** 13:15 hrs
(c)General Diary Reference (रोजानामचा संदर्भ): **Entry No.(प्रविष्टि सं.):** 026A **Date/Time(दिनांक/समय):** 15/01/2016 13:15

4. Type of Information (सूचना का प्रकार): Written

5. Place of Occurrence (घटनास्थल):
(a) Direction and Distance from P.S (थाना से दूरी और दिशा): , **Beat No(बीट सं.) :**
(b) Address(पता):
(c) In case, Outside the limit of the Police Station (यदि थाना सीमा के बाहर हैं):
Name of P.S(थाना का नाम): **District(ज़िला):**

6. Complainant / Informant (शिकायतकर्ता/सूचनाकर्ता):
(a)Name(नाम): SMT KISHNI DEVI (W/O) Shri Jagdish Chander
(b)Date/Year of Birth (जन्म तिथि /वर्ष): **Nationality (राष्ट्रीयता):** INDIA
(c)Passport No.(पासपोर्ट सं.): **Date of Issue (जारी करने की तिथि):** **Place of Issue (जारी करने का स्थान):**
(d)Occupation (व्यवसाय):
(e)Address(पता): D-168/A-2, FREEDOM FIGHTER COLONY ND, SOUTH, DELHI, INDIA,

7. Details of Known/Suspect/Unknown accused with full particulars(attach separate sheet if necessary)(ज्ञात/ संदिग्ध /अज्ञात अभियुक्त का का पुरे विवरण सहित वर्णन):

8. Reason for delay in reporting by the complainant/informant (शिकायतकर्ता / सूचनाकर्ता द्वारा रिपोर्ट देरी से दर्ज कराने के कारण):
NO DELAY

9. Particulars of the properties stolen/involved (attach separate sheet if necessary):
Sl.No. (क्र.सं.) Property Type(Description) **Est. Value(Rs.)(मूल्य (रु में))**

10. Total value of property stolen (चोरी हुई सम्पत्ति का कुल मूल्य):
11. Inquest Report / U.D. Case No., if any (मृत्यु समीक्षा रिपोर्ट / यू.डी.प्रकरण न., यदि कोई हो):

1

District : SOUTH P.S: MEHRAULI Year: 2016 FIR No: 0136 Date: 15/01/2016

12. F.I.R. Contents (attach separate sheet, if required)(प्रथम सूचना रिपोर्ट तथ्य):

In The Court of Cheif Metropolitan Magistrate District South, Saket Courts New Delhi Complaint Case of 2014. In the Matter of Smt. Kishni Devi, w/o Shri Jagdish Chander, r/o D-168/A-2, Freedom fighter Enclave IGNOU Road, Neb Sarai New Delhi-110068 Complainant versus 1. Sayed Afaq Meraj @ Afaq Miraj Syed @ Afaq s/o Shri Syed Mehrajuddin r/o- D-162, LG, Freedom Fighter Enclave Neb Sarai, New Delhi 110068, Also at C-67 Upper Ground floor (UGF) Freedom Fighter Enclaves, Neb Sarai New Delhi, 2. officer/Cleark MTNL, Hauz Khas New Delhi, 3. Area Line man, MTNL Neb Sarai Delhi, 4. SHO PS Mehrauli New Delhi Proforma Party PS Mehrauli South District South. Complaint u/s 200 CrPC for taking cognizance against the accused persons under section 417/420/442/452/464/467/468/471/506 of IPC. The complainant most respectfully submits as under. 1. That the complainant is a law abiding and peace loving citizen of India and is residing with his family member at the address as mentioned above. 2. That the above named accused persons, acting in collusion with each other, have dishonestly managed/arranged to get a new telephone connection of MTNL land line (No..29536247) is installed on the complainant's address, with the help of false and fabricated documents, by stealing the election card and photograph of the complainant from her house and by forging her signature with the intention of causing wrongful loss to the complainant and her family members and now the accused no 1 is passing threats of dire consequences. 3. That a deal related to the property of the complainant was fixed between the complainant and the accused person on 15.03.2013 and an agreement to sell was executed for a total amount of consideration of Rs 1,75,00,000/- wherein an amount of Rs 10,00,000/- was paid to the complainant by the accused person as earnest money and rest of the money was settled to be paid in parts during the period of eight months and as per the condition of the agreement to sell, original documents related to the property were handed over to the accused persons, beleiving in good faith on the representation made by him, that he get the documents/title verified through various departments. 4. That on 03.5.2013 the accused person approached the complainat and said that the deal fixed between them on 15.03.2013 related to the property was not as per the market value of the property and hence after deliberate discussions and relying on the representation made by the accused persons, the complainant agreed to execute a new advance receipt cum agreement to sell and purchase for a total consideration of Rs 1,40,00,000/- with a lesser value of Rs 35,00,000/- the the original amount fixed in previous agreement dated 15.03.2013 and the earnest amount of Rs 10,00,000/- paid at the time of previous agreement was adjusted in the new agreement and it was decided that the balance amount of Rs. 1,30,00,000/- shall be paid in parts in a period of eight months and after the completion of eight months the complainant shall handover peaceful possession and execute the necessary documents in the nameof accused person. 5. That on 24.10.2013 the complainant received a telephone bill from MTNL at the residence of the complainant, which was on the name of the accused person and was showing the address of the complainant. The husband of the complainant went to the office of MTNL to get details, as why and how the address of the complainant came on the bill of the said connection, which is not onthe complainant's name and to get the particulars stated in the bill verified. It was revealed by one of the officer of the MTNL that the said connection is on the name of the above named accsued person and it was also disclosed by that officer that if any new connection is installed then for that purpose, proper address proof and photo ID is required and if the connection is installed in a premised were the complainant does not have proper address proof then in that case the NOC of the owner or landlord is taken on record and after doing the verification of the address new connection is installed, It is pertinent to mention here that all these work of verification and installation of the connection is done at the site by the employees/staff of the MTNL, were the telephone connection is required as per the applicant demand. 6. That as per the requirement of the MTNL the complainant never applied for any

2

District : SOUTH P.S: MEHRAULI Year: 2016 FIR No: 0136 Date: 15/01/2016

connection neither the complainant made any request of any kind, nor she gave any NOC to any person for the same. It is pertinent to mention here that no address verification was conducted at the residential address of the complainant no employee/representative/staff visited at the resident of the complainant for the purpose of installation of any connection of MTNL without following the prescribed procedure. 7., That after getting all the details from the MTNL, the complainant approached the above named accused person, for the purpose of getting the clarification that why and how he managed to get this false connection installed on his name, then he started that one day in the recent past when he visited at the residence of the complainant in pretext of discussing the details of the deal of the complainant's property fixed with im and taking advantage of the fact that the complainant was alone at home, he took the complainant's election card and the passport size photograph which were lying on the table and concealed them in his pocket the complainant requested him to return the election card and photographs but he refused and said that he will use the same for the vaious other purposes and will fabricate many other documents on his name by forging the signatures of the complainant and will cause wrongful loss to the complainant and her family members. 8. That on 01.11.2013 a complaint was filed with PS Mehrauli, New Delhi stating the fact of cheating, forgery, fabricating of false documents and stealing of documents from the house of the complainant vide DD No 59B, dated 01.11.2013. the copy of the same was alos sent to D.C.P Office, Delhi on 01.11.2013 further the rectified complaint was also filed with Commissioner of Police and DCP by speed post on 07.11.2013. 9. That on 04.01.2013 at around 12.00 PM the complainat was alone at her home as it was a day of festival and her family members had gone to temple to unknown persons wearing skullcaps entered in the house of the complainant with a long knife in their hand and started using filthy language for the complainant and threatened that if the complainant will not vacate her house then she shall remain ready to face dire consequences. This fact of passing threats was also brought the the notice of SHO, PS Mehrauli Delhi Vide DD No 46B, Dated 04/11/2013. 11. That the SHO in the present matter has refused to register the FIR and also has refused to recover the election card and photographs of the complainant from the accused No 1, it was also requested by the complainant to SHO to recover the false and fabricated documents from MTNL Hauz Khas but the SHO refused to take any action against the accused person. 12. That it seems that the police is also in favour of the accused persons and is not going to take any action against them. Feeling aggrieved by the inaction on the part of the police personnel the complainant is left with no other efficacious remedy except to file the present complaint. 13. That the unlawful act committed by the accused persons comes under the category of cognizable offence. 14. that the offence is committed within the jurisdiction of this Hon'ble court. Prayer it is therefore most respectfully prayed that this Hon'ble court may be pleased. a. To take cognzance against the accused No.1 to 3 u/s 417/420/442/452/463/464/467/468/471/506 of IPC and issue appropriate process against the Accused persons, try and punish them according to law and in the interest of Justice. b. Pass such other order(s) as deem fit and proper in the facts of the present case. Duty officer PS Mehrauli the present complaint is pending u/s 156.3 Cr.P.C in the court of Ms Shreya Arora Mehta MM Saket Court New Delhi. Enquiry has been conducted. From the contents of the complaint the enquiry conducted so far offence u/s 451/506/509/34 IPC is made out. As far as other allegation are concerned, the enquiry is going on and relevant sections of law shall be added in this case if required. You are directed to register a case and mark investigation to me by order SHO/Mehrauli. Date and time of handing over tehreer- 15/1/2016 at 1.10 PM. SD English SI Manoj Singh PS Mehrauli New Delhi Dt 15/1/16.कार्यवाही पुलिस अज थाना तहरीर की दर पेशगी पर मुकदमा हजा बाजुर्म मजकूर u/s 451/506/509/34 IPC दर्ज कम्प्युटर किया जाकर असल तहरीर मय Computer Copy of FIR बराये आईन्दा तफ्तीश बाहुक्म जनाब SHO साहब हवाले SI Manoj Singh साहब की जा रही है। दीगर नकुलात बजरिये डाक अफसरान बाला की खिदमत में अरसाल होगें। HC/DO.

3

District : SOUTH P.S: MEHRAULI Year: 2016 FIR No: 0136 Date: 15/01/2016

13. Action Taken Since the above information reveals commission of offence(s) u/s as mentioned at Item No. 2:
(की गयी कार्यवाही: चूंकि उपरोक्त जानकारी से पता चलता है कि किया गया अपराध मद सं.2.में उल्लेख धारा के तहत है):
(i) Registered the case and took up the investigation: **OR (या)**
(प्रकरण दर्ज किया गया और जांच के लिए लिया गया):
(ii) Directed (Name of the I.O.)(जांच अधिकारी का नाम): MANOJ SINGH CHUPHAL **Rank (पद):** SI (SUB-INSPECTOR)
No(सं.): 16100158 **to take up the investigation (को जांच आपने पास में लेने के लिए निर्देश दिया गया) OR(या)**
(iii) Refused investigation due to(जांच के लिए): **OR (के कारण इंकार किया या)**
(iv) Transferred to P.S(name)(थाना): **District(ज़िला):**
on point of jurisdiction (को क्षेत्राधिकार के कारण हस्तांतरित)

F.I.R read over to the complainant/informant,admitted to be correctly recorded and a copy given to the complainant/informant, free of cost : (शिकायतकर्ता / सूचनाकर्ता को प्राथमिकी पढ़ कर सुनाई गयी, सही दर्ज हुई माना और एक कॉपी निशुल्क शिकायतकर्ता को दी गयी) :

R.O.A.C.(आर.ओ.ए.सी.):

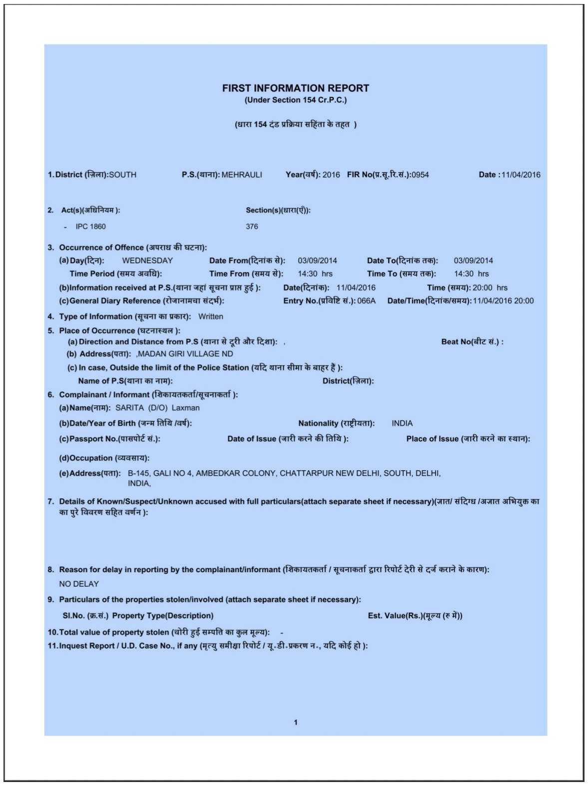
14. Signature / Thumb Impression of the Complainant / Informant:
(शिकायतकर्ता / सूचनाकर्ता के हस्ताक्षर / अंगूठे का निशान):

Signature of Officer
Name(नाम): SATISH KUMAR
Rank (पद): HC (HEAD CONSTABLE)
No.(सं.): 28821146

15. Date and Time of despatch to the court:
(अदालत में प्रेषण की दिनांक और समय):

4

Figure A2: Snapshot of challenges faced during parsing the data into Excel



FIRST INFORMATION REPORT
(Under Section 154 Cr.P.C.)

(धारा 154 दंड प्रक्रिया सहिंता के तहत)

1. District (ज़िला):SOUTH P.S.(थाना): MEHRAULI Year(वर्ष): 2016 FIR No(प्र.सू.रि.सं.):0954 Date : 11/04/2016

2. Act(s)(अधिनियम): Section(s)(धारा(एँ)):
- IPC 1860 376

3. Occurrence of Offence (अपराध की घटना):
(a) Day(दिन): WEDNESDAY Date From(दिनांक से): 03/09/2014 Date To(दिनांक तक): 03/09/2014
Time Period (समय अवधि): Time From (समय से): 14:30 hrs Time To (समय तक): 14:30 hrs
(b)Information received at P.S.(थाना जहां सूचना प्राप्त हुई): Date(दिनांक): 11/04/2016 Time (समय): 20:00 hrs
(c) General Diary Reference (रोजानामचा संदर्भ): Entry No.(प्रविष्टि सं.): 066A Date/Time(दिनांक/समय): 11/04/2016 20:00

4. Type of Information (सूचना का प्रकार): Written

5. Place of Occurrence (घटनास्थल):
(a) Direction and Distance from P.S (थाना से दूरी और दिशा): , Beat No(बीट सं.) :
(b) Address(पता): ,MADAN GIRI VILLAGE ND
(c) In case, Outside the limit of the Police Station (यदि थाना सीमा के बाहर हैं):
Name of P.S(थाना का नाम): District(ज़िला):

6. Complainant / Informant (शिकायतकर्ता/सूचनाकर्ता):
(a) Name(नाम): SARITA (D/O) Laxman
(b)Date/Year of Birth (जन्म तिथि /वर्ष): Nationality (राष्ट्रीयता): INDIA
(c) Passport No.(पासपोर्ट सं.): Date of Issue (जारी करने की तिथि): Place of Issue (जारी करने का स्थान):
(d)Occupation (व्यवसाय):
(e) Address(पता): B-145, GALI NO 4, AMBEDKAR COLONY, CHATTARPUR NEW DELHI, SOUTH, DELHI, INDIA,

7. Details of Known/Suspect/Unknown accused with full particulars(attach separate sheet if necessary)(ज्ञात/ संदिग्ध /अज्ञात अभियुक्त का का पुरे विवरण सहित वर्णन):

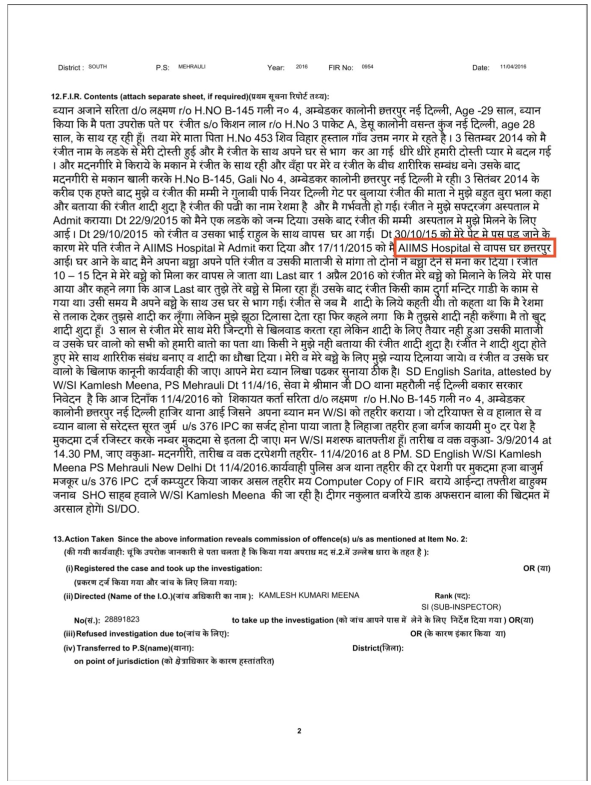
8. Reason for delay in reporting by the complainant/informant (शिकायतकर्ता / सूचनाकर्ता द्वारा रिपोर्ट देरी से दर्ज कराने के कारण):
NO DELAY

9. Particulars of the properties stolen/involved (attach separate sheet if necessary):
Sl.No. (क्र.सं.) Property Type(Description) Est. Value(Rs.)(मूल्य (रु में))

10. Total value of property stolen (चोरी हुई सम्पत्ति का कुल मूल्य): -

11. Inquest Report / U.D. Case No., if any (मृत्यु समीक्षा रिपोर्ट / यू.डी.प्रकरण न., यदि कोई हो):

1



District : SOUTH P.S: MEHRAULI Year: 2016 FIR No: 0954 Date: 11/04/2016

12. F.I.R. Contents (attach separate sheet, if required)(प्रथम सूचना रिपोर्ट तथ्य):
ब्यान अजाने सरिता d/o लक्ष्मण r/o H.NO B-145 गली न० 4, अम्बेडकर कालोनी छत्तरपुर नई दिल्ली, Age -29 साल, ब्यान किया कि मै पता उपरोक्त पते पर रंजीत s/o किशन लाल r/o H.No 3 पाकेट A, डेसू कालोनी वसन्त कुंज नई दिल्ली, age 28 साल, के साथ रह रही हूँ। तथा मेरे माता पिता H.No 453 शिव विहार हस्ताल गाँव उत्तम नगर मे रहते है। 3 सितम्बर 2014 को मै रंजीत नाम के लडके से मेरी दोस्ती हुई और मै रंजीत के साथ अपने घर से भाग कर आ गई धीरे धीरे हमारी दोस्ती प्यार मे बदल गई। और मदनगीरि मे किराये के मकान मे रंजीत के साथ रही और वँहा पर मेरे व रंजीत के बीच शारीरिक सम्बंध बने। उसके बाद मदनगीरी से मकान खाली करके H.No B-145, Gali No 4, अम्बेडकर कालोनी छत्तरपुर नई दिल्ली मे रही। 3 सितंबर 2014 के करीब एक हफ्ते बाद मुझे व रंजीत की मम्मी ने गुलाबी पार्क नियर दिल्ली गेट पर बुलाया रंजीत की माता ने मुझे बहुत बुरा भला कहा और बताया की रंजीत शादी शुदा है रंजीत की पत्नी का नाम रेशमा है और मै गर्भवती हो गई। रंजीत ने मुझे सफ्दरजंग अस्पताल मे Admit कराया। Dt 22/9/2015 को मैने एक लडके को जन्म दिया। उसके बाद रंजीत की मम्मी अस्पताल मे मुझे मिलने के लिए आई। Dt 29/10/2015 को रंजीत व उसका भाई राहुल के साथ वापस घर आ गई। Dt 30/10/15 को मेरे पेट मे पस पड जाने के कारण मेरे पति रंजीत ने AIIMS Hospital मे Admit करा दिया और 17/11/2015 को मै AIIMS Hospital से वापस घर छत्तरपुर आई। घर आने के बाद मैने अपना बच्चा अपने पति रंजीत व उसकी माताजी से मांगा तो दोनो ने बच्चा देने से मना कर दिया। रंजीत 10 – 15 दिन मे मेरे बच्चे को मिला कर वापस ले जाता था। Last बार 1 अप्रैल 2016 को रंजीत मेरे बच्चे को मिलाने के लिये मेरे पास आया और कहने लगा कि आज Last बार तुझे तेरे बच्चे से मिला रहा हूँ। उसके बाद रंजीत किसी काम दुर्गा मन्दिर गाडी के काम से गया था। उसी समय मै अपने बच्चे के साथ उस घर से भाग गई। रंजीत से जब मै शादी के लिये कहती थी। तो कहता था कि मै रेशमा से तलाक देकर तुझसे शादी कर लूँगा। लेकिन मुझे झूठा दिलासा देता रहा फिर कहले लगा कि मै तुझसे शादी नही करुँगा। मै तो खुद शादी शुदा हूँ। 3 साल से रंजीत मेरे साथ मेरी जिन्दगी से खिलवाड करता रहा लेकिन शादी के लिए तैयार नही हुआ उसकी माताजी व उसके घर वालो को सभी को हमारी बातो का पता था। किसी ने मुझे नही बताया की रंजीत शादी शुदा है। रंजीत ने शादी शुदा होते हुए मेरे साथ शारिरीक संबंध बनाए व शादी का धौखा दिया। मेरी व मेरे बच्चे के लिए मुझे न्याय दिलाया जाये। व रंजीत व उसके घर वालो के खिलाफ कानूनी कार्यवाही की जाए। आपने मेरा ब्यान लिखा पढकर सुनाया ठीक है। SD English Sarita, attested by W/SI Kamlesh Meena, PS Mehrauli Dt 11/4/16, सेवा मे श्रीमान जी DO थाना महरौली नई दिल्ली बकार सरकार निवेदन है कि आज दिनाँक 11/4/2016 को शिकायत कर्ता सरिता d/o लक्ष्मण r/o H.No B-145 गली न० 4, अम्बेडकर कालोनी छत्तरपुर नई दिल्ली हाजिर थाना आई जिसने अपना ब्यान मन W/SI को तहरीर कराया। जो दरियाफ्त से व हालात से व ब्यान बाला से सरेदस्त सूरत जुर्म u/s 376 IPC का सर्जद होना पाया जाता है लिहाजा तहरीर हजा बर्गज कायमी मु० दर पेश है मुकदमा दर्ज रजिस्टर करके नम्बर मुकदमा से इतला दी जाए। मन W/SI मशरुफ बातफ्तीश हूँ। तारीख व वक्त वकुआ- 3/9/2014 at 14.30 PM, जाए वकुआ- मदनगीरी, तारीख व वक्त दरपेशगी तहरीर- 11/4/2016 at 8 PM. SD English W/SI Kamlesh Meena PS Mehrauli New Delhi Dt 11/4/2016.कार्यवाही पुलिस अज थाना तहरीर की दर पेशगी पर मुकदमा हजा बाजुर्म मजकूर u/s 376 IPC दर्ज कम्प्युटर किया जाकर असल तहरीर मय Computer Copy of FIR बराये आईन्दा तफ्तीश बाहुक्म जनाब SHO साहब हवाले W/SI Kamlesh Meena की जा रही है। दीगर नकुलात बजरिये डाक अफसरान बाला की खिदमत में अरसाल होगें। SI/DO.

13. Action Taken Since the above information reveals commission of offence(s) u/s as mentioned at Item No. 2:
(की गयी कार्यवाही: चूंकि उपरोक्त जानकारी से पता चलता है कि किया गया अपराध मद सं.2.में उल्लेख धारा के तहत है):
(i) Registered the case and took up the investigation: OR (या)
(प्रकरण दर्ज किया गया और जांच के लिए लिया गया):
(ii) Directed (Name of the I.O.)(जांच अधिकारी का नाम): KAMLESH KUMARI MEENA Rank (पद): SI (SUB-INSPECTOR)
No(सं.): 28891823 to take up the investigation (को जांच आपने पास में लेने के लिए निर्देश दिया गया) OR(या)
(iii) Refused investigation due to(जांच के लिए): OR (के कारण इंकार किया या)
(iv) Transferred to P.S(name)(थाना): District(ज़िला):
on point of jurisdiction (को क्षेत्राधिकार के कारण हस्तांतरित)

2

Subsequently, I tackled the challenge of converting these unstructured scanned PDFs, which are essentially images of filed reports, into a structured Excel format for further processing. Compounding this challenge was the bilingual nature of the FIRs, featuring text in both English and Hindi. Figure A2 shows a snapshot of the scanned nature of the PDF and the two languages co-occurring. Addressing these issues, I harnessed *Optical Character Recognition* (OCR) techniques, leveraging tools such as *pytesseract*, *BeautifulSoup*, and *pdf2image*. This approach involved strategically creating boxes around the relevant content in the scanned PDFs, ultimately allowing for the extraction and parsing of data into Excel. Figure A3 visually represents the creation of these boxes, while Figure A4 offers a snapshot of the organized Excel data.

While these strides represent significant progress, the ongoing task involves identifying and classifying reported crimes specifically targeted against women. To achieve this, a Machine Learning algorithm is being deployed to comprehend the crime descriptions and discern

Figure A3: Snapshot of boxes around the text


FIRST INFORMATION REPORT
(Under Section 154 Cr.P.C.)

(धारा 154 दंड प्रक्रिया सहिंता के तहत)

1. District (ज़िला): Distric_1 WEST P.S.(थाना): PS_1 KIRTI NAGAR Year(वर्ष): YEAR_1 2016 FIR No(प्र.सू.रि.सं.): FIR_1 0153 Date : Date_1 09/03/2016

2. Act(s)(अधिनियम): Acts_2 - IPC 1860 Section(s)(धारा(एँ)): Sections_2 356/379/34

3. Occurrence of Offence (अपराध की घटना):
(a) Day(दिन): A_Day_3 WEDNESDAY Date From(दिनांक से): A_DateFrom_3 09/03/2016 Date To(दिनांक तक): A_DateTo_3 09/03/2016
Time Period (समय अवधि): A_TimePeriod_3 Time From (समय से): A_TimeFrom_3 21:30 hrs Time To (समय तक): A_TimeTo_3 21:30 hrs
(b) Information received at P.S.(थाना जहां सूचना प्राप्त हुई): B_Information_3 Date(दिनांक): B_Date_3 09/03/2016 Time (समय): B_Time_3 23:00 hrs
(c) General Diary Reference (रोजानामचा संदर्भ): C_General_3 Entry No.(प्रविष्टि सं.): C_Entry_3 042A Date/Time(दिनांक/समय): C_Date_3 09/03/2016 23:20

4. Type of Information (सूचना का प्रकार): TypeInfo_4 Written

5. Place of Occurrence (घटनास्थल):
(a) Direction and Distance from P.S (थाना से दूरी और दिशा): A_Direction_5 EAST , 1.5 KILOMETERS Beat No(बीट सं.) : A_Beat_5 8
(b) Address(पता): B_Address_5 IN FRONT OF H NO. A-2 ON ROAD ,OLD QUATER ,RAMESH NAGAR ,NEW DELHI
(c) In case, Outside the limit of the Police Station (यदि थाना सीमा के बाहर हैं):
Name of P.S(थाना का नाम): C_InCaseName_5 District(ज़िला): C_InCaseDistrict_5

6. Complainant / Informant (शिकायतकर्ता/सूचनाकर्ता):
(a) Name(नाम): A_Name_6 MS AKANKSHA AGGARWAL (D/O) Sh Rajeev Ranjan
(b) Date/Year of Birth (जन्म तिथि /वर्ष): B_Date_6 1990 Nationality (राष्ट्रीयता): B_Nationality_6 INDIA
(c) Passport No.(पासपोर्ट सं.): C_Passaport_6 Date of Issue (जारी करने की तिथि): C_Date_6 Place of Issue (जारी करने का स्थान): C_Place_6
(d) Occupation (व्यवसाय): D_Occupation_6
(e) Address(पता): E_Address_6 H NO. A-2, OLD QUARTER, RAMESH NAGAR, NEW DELHI, KIRTI NAGAR, WEST, DELHI, 110015, INDIA, 8826894779,

7. Details of Known/Suspect/Unknown accused with full particulars(attach separate sheet if necessary)(ज्ञात/ संदिग्ध /अज्ञात अभियुक्त का का पुरे विवरण सहित वर्णन):
[8] in page 1

8. Reason for delay in reporting by the complainant/informant (शिकायतकर्ता / सूचनाकर्ता द्वारा रिपोर्ट देरी से दर्ज कराने के कारण):
Reason_8 NO DELAY

9. Particulars of the properties stolen/involved (attach separate sheet if necessary):
Sl.No. (क्र.सं.) Property Type(Description) Est. Value(Rs.)(मूल्य (रु में))
SINO_9 PropertyType_9 EstValue_9

10. Total value of property stolen (चोरी हुई सम्पत्ति का कुल मूल्य): [10] in page 1 10000

11. Inquest Report / U.D. Case No., if any (मृत्यु समीक्षा रिपोर्ट / यू.डी.प्रकरण न., यदि कोई हो):
[12] in page 1

1

both the victim and the accused parties. Figure A5 shows the names of the accused and victim from one of the FIRs. This phase presents an added complication due to the bilingual aspect of the FIRs. My proposed strategy to surmount this challenge entails several steps. Currently, I have utilized the ‘Section(s)’ field, which denotes the pertinent section of the Indian Penal Code 1860 that each crime corresponds to. Although this filtering method will not capture all reported crimes against women, it serves as an initial screening process.

Figure A4: Snapshot of data parsed into Excel

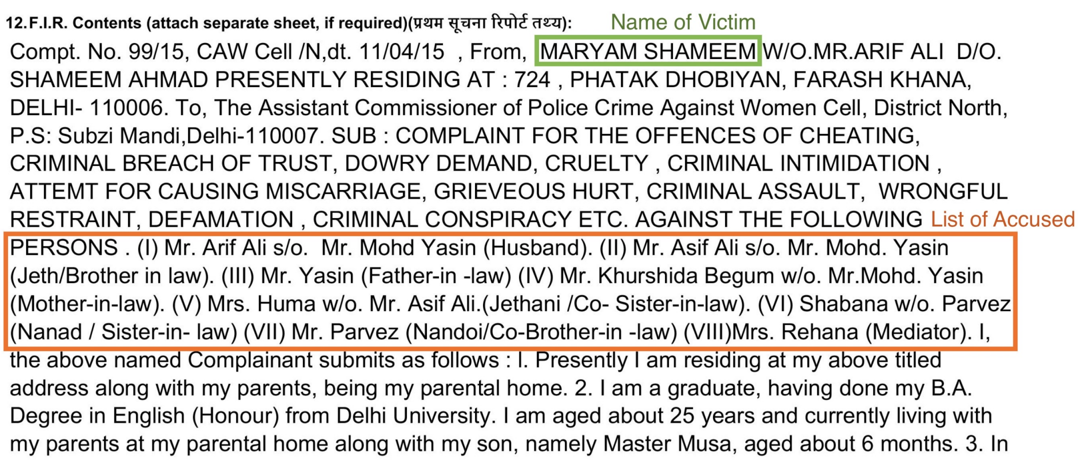
	A	B	C	D	E	F	G	H
1		1				R	F	G
2		District	P.S.	Year	FIR No	Date	Act(s)	Section(s)
3								
4	763	METRO	161 AIRPORT METRO	2019	11	24/09/201	IPC 1860 N	379/511/3
5	891	METRO	Metro Police Station Azadpur	2019	6	18/10/201	- IPC 1860	379
6	863	METRO	KASHMIRI GATE METRO	2019	78	31/10/201	- IPC 1860	379/511
7	907	METRO	Metro Police Station Ghitorni	2019	11	23/08/201	- IPC 1860	379/511
8	967	METRO	Metro Police Station Pragati Ram? al a tam	2019	7	13/04/201	- IPC 1860	379
9	103	METRO	RAJA GARDEN METRO	2019	2	15/01/201	- IPC 1860	323/341
10	866	METRO	KASHMIRI GATE METRO	2019	81	11/11/201	- IPC 1860	379/511
11	836	METRO	KASHMIRI GATE METRO	2019	51	15/07/201	- IPC 1860	379/411
12	978	METRO	Metro Police Station Pragati Ram? al a tam	2019	18	19/11/201	IPC 1860 T	509/354(A)
13	777	METRO	JANAK PURI METRO	2019	3	15/03/201	- IPC 1860	354/354(D
14	107	METRO	SHASHTRI PARK METRO	2019	21	24/09/201	- IPC 1860	365
15	840	METRO	KASHMIRI GATE METRO	2019	55	30/07/201	- IPC 1860	379/511
16	794	METRO	KASHMIRI GATE METRO	2019	9	12/02/201	- IPC 1860	379
17	106	METRO	SHASHTRI PARK METRO	2019	9	10/03/201	- IPC 1860	384
18	111	METRO	YAMUNA DEPOT METRO	2019	27	06/12/201	- IPC 1860	379/411
19	818	METRO	KASHMIRI GATE METRO	2019	33	08/05/201	- IPC 1860	379
20	103	METRO	RAJA GARDEN METRO	2019	4	17/01/201	- IPC 1860	379
21	808	METRO	KASHMIRI GATE METRO	2019	23	31/03/201	- IPC 1860	379/411

	A	AP	AQ	AR	AS
1		12	Action taken since the above information reveals commission of offe		
2		F.I.R. Contents	e and took u	Directed (Name of the I.O.)	Rank
3					
4	763	ब्यान अजाने Rajesh Kakkar En		SUNITA DUTTA	SI (SUB-INSPECTOR)
5	891	ब्यान अजाने दीपक कुमार S/o		SURESH KUMAR	ASST. SI (ASSISTANT SUB-INSPECTOR)
6	863	ब्यान अजाने श्री शाह आलम S/		TILAK RAJ	HC (HEAD CONSTABLE)
7	907	ब्यान अजाने Abhay PratapS/o		RANBIR SINGH	SI (SUB-INSPECTOR)
8	967	श्रीमान SHO PS PRAGATI MAID		UDAI PAL SINGH	ASST. SI (ASSISTANT SUB-INSPECTOR)
9	103	सेवा मे, श्री मान SHO साहब , पु		ANITA KUMARI	ASST. SI (ASSISTANT SUB-INSPECTOR)
10	866	ब्यान अजाने Sh Amrinder Sing		CHANDRA SHEKHAR	ASST. SI (ASSISTANT SUB-INSPECTOR)
11	836	ब्यान अजाने रवि सिंह S/o दान		RAJESH KUMAR	ASST. SI (ASSISTANT SUB-INSPECTOR)
12	978	सेवा में श्रीमान जी SHO, Thana		BABBAR BHAN	| INSPECTOR)
13	777	SIM अजाने मोनिका D/O रामसे		RAJBIR SINGH	ASST. SI (ASSISTANT SUB-INSPECTOR)
14	107	DD No.21A Dt.20/10/14 P.S. S		VIJAY SINGH	ASST. SI (ASSISTANT SUB-INSPECTOR)
15	840	ब्यान अजाने निर्मल कुमार शर्मा		NEERAJ GAUTAM	HC (HEAD CONSTABLE)
16	794	सेवा मे, S.HO Kashmiri Gate M		NEERAJ GAUTAM	HC (HEAD CONSTABLE)
17	106	ब्यान अजाने रंजीत S/O पंचमल		VIJAY SINGH	ASST. SI (ASSISTANT SUB-INSPECTOR)
18	111	सेवा मे श्रीमान जी SHO यमुना		VIRENDRA SINGH YADAV	ASST. SI (ASSISTANT SUB-INSPECTOR)
19	818	To, SHO K. Gate Metro Delhi,		SANJEEV KUMAR	HC (HEAD CONSTABLE)
20	103	सेवामे, श्रीमान SHO साहब पुलि		HARPAL SINGH	ASST. SI (ASSISTANT SUB-INSPECTOR)
21	808	ब्यान अजाने अनीश कुमार S/o		NEERAJ GAUTAM	HC (HEAD CONSTABLE)

To enhance accuracy, I intend to incorporate the details of the complainant/informant and suspects to refine the data further. Employing *Named Entity Recognition* (NER), I have extracted all names from the crime descriptions. All the names, except that of the victim, belong to other columns - police station, officer, and accused names. This exercise has given me the names of the victim and the accused as separate columns, thereby forming the basis of my training data set. Subsequently, I have trained a model on this data set and evaluated its performance on the test data set, which comprises the remaining FIRs. Presently, I am in the process of checking the accuracy of this exercise.

Finally, I will determine the gender of the victim and the accused by analyzing phrases such as ‘W/O’ and ‘D/O’ after a name (indicating female) or ‘S/O’ (indicating male) and augment the gender classification using the *NamSor* package in R. This approach represents a refinement compared to the *NamSor* package as its accuracy to detect gender with names written in Hindi language is 45-55%.[48]

Figure A5: Snapshot of extraction of names of victim and accused



12.F.I.R. Contents (attach separate sheet, if required)(प्रथम सूचना रिपोर्ट तथ्य):

Compt. No. 99/15, CAW Cell /N,dt. 11/04/15 , From, MARYAM SHAMEEM W/O.MR.ARIF ALI D/O. SHAMEEM AHMAD PRESENTLY RESIDING AT : 724 , PHATAK DHOBIYAN, FARASH KHANA, DELHI- 110006. To, The Assistant Commissioner of Police Crime Against Women Cell, District North, P.S: Subzi Mandi,Delhi-110007. SUB : COMPLAINT FOR THE OFFENCES OF CHEATING, CRIMINAL BREACH OF TRUST, DOWRY DEMAND, CRUELTY , CRIMINAL INTIMIDATION , ATTEMT FOR CAUSING MISCARRIAGE, GRIEVEOUS HURT, CRIMINAL ASSAULT, WRONGFUL RESTRAINT, DEFAMATION , CRIMINAL CONSPIRACY ETC. AGAINST THE FOLLOWING PERSONS . (I) Mr. Arif Ali s/o. Mr. Mohd Yasin (Husband). (II) Mr. Asif Ali s/o. Mr. Mohd. Yasin (Jeth/Brother in law). (III) Mr. Yasin (Father-in -law) (IV) Mr. Khurshida Begum w/o. Mr.Mohd. Yasin (Mother-in-law). (V) Mrs. Huma w/o. Mr. Asif Ali.(Jethani /Co- Sister-in-law). (VI) Shabana w/o. Parvez (Nanad / Sister-in- law) (VII) Mr. Parvez (Nandoi/Co-Brother-in -law) (VIII)Mrs. Rehana (Mediator). I, the above named Complainant submits as follows : I. Presently I am residing at my above titled address along with my parents, being my parental home. 2. I am a graduate, having done my B.A. Degree in English (Honour) from Delhi University. I am aged about 25 years and currently living with my parents at my parental home along with my son, namely Master Musa, aged about 6 months. 3. In

This comprehensive strategy will culminate in a detailed, micro-level data set encompassing reported crimes against women in Delhi from 2016 to 2019.

Currently, I have focused on using the ‘Section(s)’ to identify the crime categories used as the outcome variables. Reports of domestic violence entail incidents of dowry-related

[48]https://namsor.app/features/gender-name#genderize-name-batch

torture or death, mistreatment of wives or daughters-in-law, deceitful marriages, or tragically coerced miscarriages. Reported HoS crimes against women consist of cases of obscene acts, sexual harassment, throwing of acid, stalking or insulting the modesty of a woman. Overall reported crimes against women include both reported domestic violence and reported HoS and reported cases of rape and female kidnappings. Finally, reported non-gender-specific street crimes include reports of theft, mischief, creating public nuisance, or obstructing public pathways. Table A1 lists the sections from Indian Penal Code 1860 which are used to construct these categories.

Table A1: Description of Crime Categories

Category	Sections from IPC 1860
All Crimes Against Women	294, 304B, 312, 313, 314, 354, 354A, 354B, 354C, 354D, 366, 366A, 366B, 372, 373, 375, 376, 376A, 376AB, 376B, 376C, 376D, 376DA, 376DB, 376E, 376F, 493, 496, 497, 498, 498A, 509, 376/511
Domestic Violence	304B, 305, 306, 312, 313, 314, 376B, 493, 496, 498A
HoS Crimes Against Women	294, 326A, 326B, 354, 354, 354A, 354B, 354C, 354D, 509
Non-Gender-Specific Street Crimes	283, 319, 320, 321, 322, 323, 324, 325, 326, 334, 335, 336, 337, 338, 349, 350, 351, 352, 355, 356, 378, 379, 425, 426, 427, 431, 432, 434, 510

Notes: This table describes the categories of crime used as outcome variables in the analysis and lists the corresponding sections from the Indian Penal Code 1860 used to identify crimes pertaining to them. *Source*: `https://www.indiacode.nic.in/handle/123456789/2263?sam_handle=123456789/1362`.

Figure A6: Snapshot of a reported crime occurring on bus

12.F.I.R. Contents (attach separate sheet, if required)(प्रथम सूचना रिपोर्ट तथ्य):

बयान अजाने दिलीप कुमार S/o Ram Raj Singh R/o A-598 near शीतला माता मंदिर, JJ Colony Madanpur Khadar Delhi 76 Age 37 years Mob 9311818181. बयान किया की मै पता उपरोक्त पर सपरिवार रहता हूँ, तथा मालिक बबलू @ मनोज कुमार के यहाँ RTV bus मे conductor का काम करता हूँ, जो मै इन दिनो बस No DL 1VA 9562 जो कि मदनपुर खादर व नेहरू place के बीच चलती है, मै कंडक्टर का काम कर रहा हूँ, जो आज दि० 31.08.19 को समय करीब 11/45 Am पर जब मै उपरोक्त RTV मे चालक virendra के साथ नेहरू प्लेस से चक्कर लगाकर मदनपुर खादर जा रहा था और जैसे ही बस रोड न. 13 पर, नाला रोड शाहीन बाग के पास पहुंची तो वहाँ पर एक लड़के ने बस को हाथ दिया ओर बस मे चढ़ गया और बस मे चढ़ते ही driver

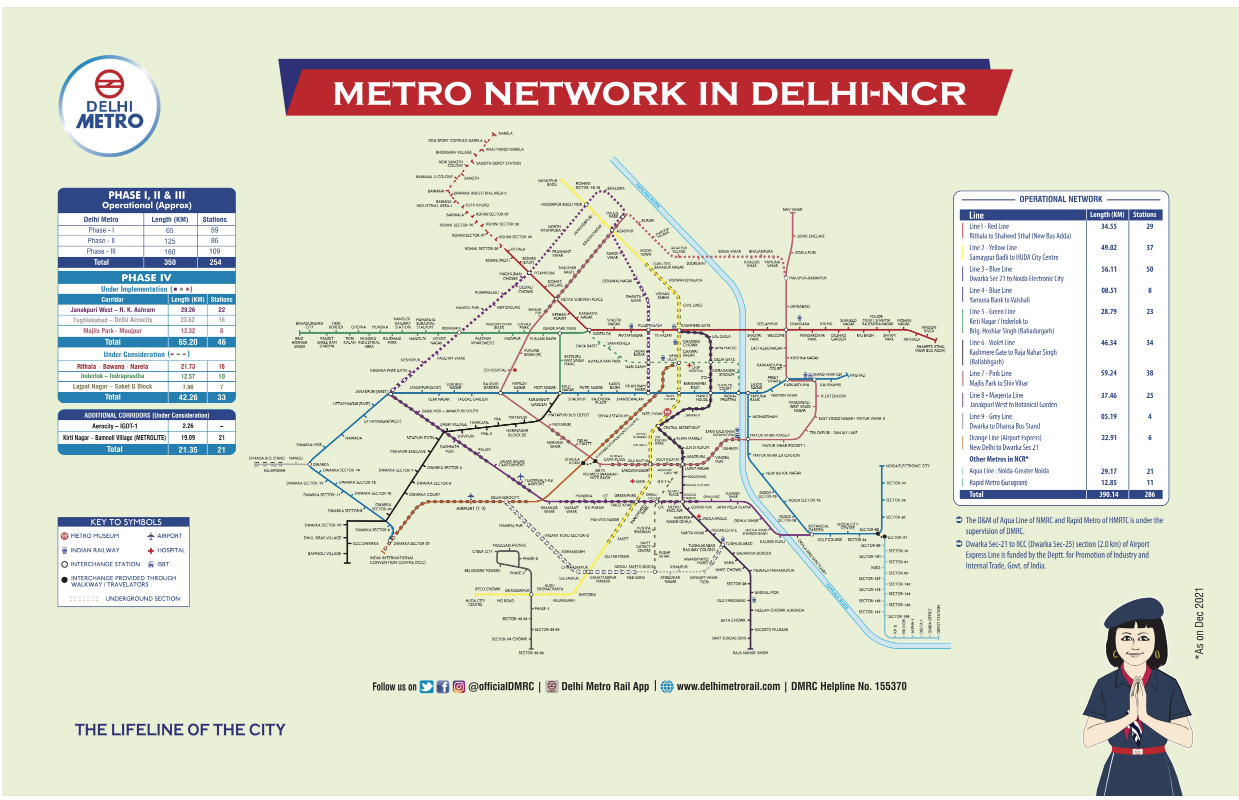
A.2 Delhi Metro - Phases and Network

Table A2: Summary statistics of the construction phases of Delhi Metro

Phases	Lines	Route	Stations	Length (in Kms)	Date
Phase-I	Red Line	Shahdara - Tis Hazari	6	8.35	25.12.2002
		Tis Hazari - Inderlok	4	4.87	04.10.2003
		Inderlok - Rithala	8	8.84	01.04.2004
	Yellow Line	Vishwavidyalaya - Kashmere Gate	4	4.06	20.12.2004
		Kashmere Gate - Central Secretariat	6	6.62	03.07.2005
	Blue Line	Barakhamba - Dwarka	22	22.74	31.12.2005
		Dwarka - Dwarka Sector 9	6	6.47	01.04.2006
		Barakhamba - Indraprastha	3	2.8	11.11.2006
Phase-II	Red Line	Shahdara – Dilshad Garden	3	3.09	03.06.2008
	Yellow Line	Vishwavidyalaya - Jahangirpuri	5	6.36	03.02.2009
		HUDA City Centre - Qutab Minar	10	15.82	21.06.2010
		Central Secretariat - Qutab Minar	9	11.76	03.09.2010
	Blue Line	Indraprastha - Yamuna Bank	1	2.17	10.05.2009
		Yamuna Bank - Noida City Centre	10	12.85	13.11.2009
		Dwarka sector 9 - Dwarka Sector 21	2	2.76	30.10.2010
		Anand Vihar ISBT - Vaishali	2.57	2	27.01.2010
	Green Line	Inderlok – Kirti Nagar - Mundka	18.46	16	
	Orange Line	New Delhi - IGI Airport(T-3) - Dwarka Sector-21	6	22.70	23.02.2011
	Violet Line	Central Secretariat - Sarita Vihar	13	15.34	03.10.2010
		Sarita Vihar - Badarpur	3	4.82	14.01.2011
Phase-III	Red Line	Dilshad Garden - Shaheed Sthal (New Bus Adda)	8	9.63	09.03.2019
	Yellow Line	Jahangirpuri - Samaypur Badli	3	4.38	10.11.2015
	Blue Line	Noida City Centre - Noida Electronic City	6	6.80	09.03.2019
	Violet Line	Central Secretariat - Janpath - Mandi House	2		28.06.2014
		ITO station			08.06.2015
		Sarai - Escorts Mujesar(Faridabad)	9	13.56	06.09.2015
		Delhi Gate – Kashmere Gate	3		28.05.2017
		Escorts Mujesar - Raja Nahar Singh (Ballabhgarh)	2	3.35	19.11.2018
	Pink Line	Majlis Park - Durgabhai Deshmukh South Campus	12	21.57	14.03.2018
		Durgabhai Deshmukh South Campus - Lajpat Nagar	6	8.53	06.08.2018
		Trilokpuri - Sanjay Lake - Shiv Vihar	15	17.86	31.10.2018
		Lajpat Nagar - Mayur Vihar Pocket-1	5	9.63	31.12.2018
	Magenta Line	Botanical Garden - Kalkaji Mandir	9	12.64	25.12.2017
		Kalkaji Mandir - Janakpuri West	16	24.82	28.05.2018
	Grey Line	Dwarka - Najafgarh	4	4.98	04.10.2019

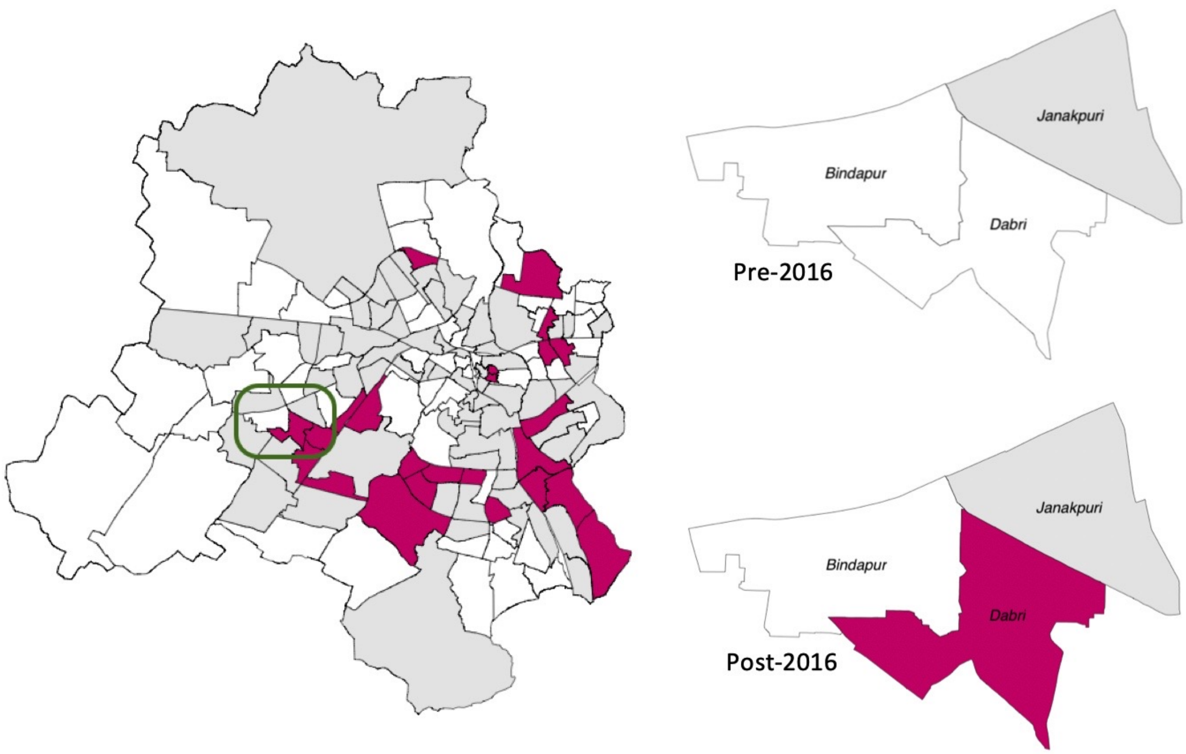
Notes: The routes highlighted in pink encompass metro stations that opened between 2016 and 2019 and serve as the first metro stations within their respective police jurisdictions. These stations will be used in the current analysis. *Source*: `https://www.delhimetrorail.com/pages/en/present-network`.

Figure A7: Metro Network in Delhi in 2021



Notes: Source:`https://www.delhimetrorail.com/pages/en/network_map`.

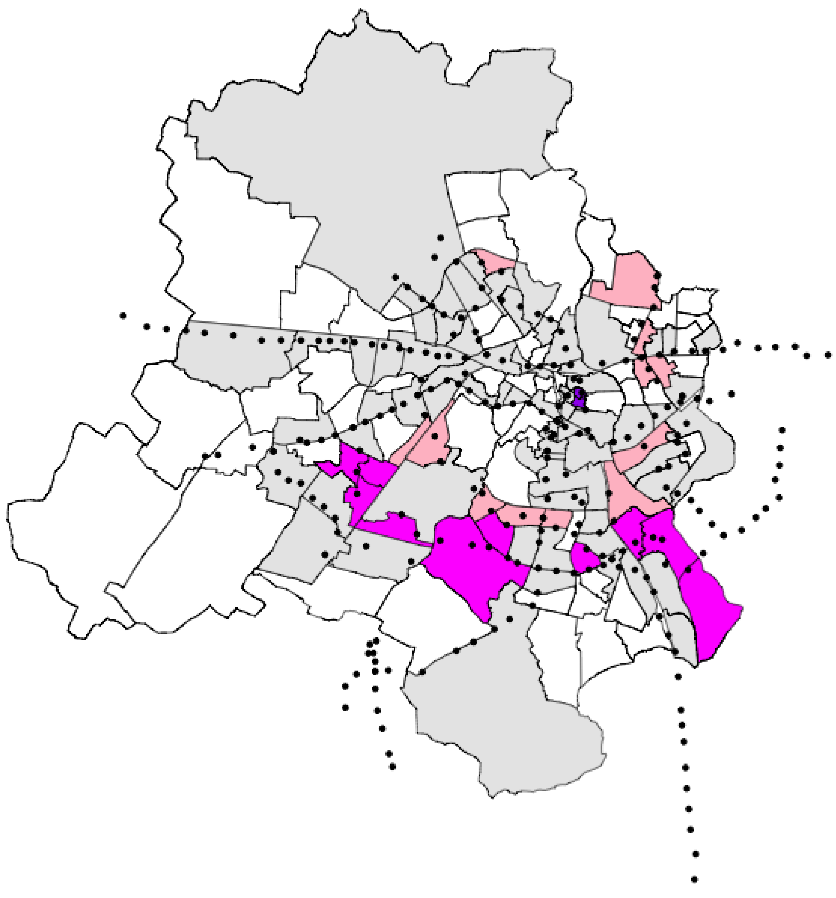
Figure A8: Treated Police Jurisdictions - An Illustration



Notes: This figure illustrates the allocation of police jurisdictions into three groups: treatment, never-treated, and already treated. The left panel displays this allocation and outlines the three jurisdictions used as an example in the text in green. The upper right panel depicts the pre-treatment period, when the police jurisdiction of *Janakpuri* had a metro station, but that of *Dabri* and *Bindapur* did not. The lower panel shows the post-treatment period when the first metro station opened in *Dabri*. Consequently, the police jurisdiction of *Dabri* falls in the treatment group, while that of *Bindapur* and *Janakpuri* remain in the control group and the already treated group, respectively.

A.3 Treatment and Control Groups

Figure A9: Treated Police Jurisdictions by Metro Line Groups



Notes: Magenta: Police jurisdictions where the first metro station that opened in the treatment period between 2016 and 2019 was on the magenta line. *Pink*: Police jurisdictions where the first metro station that opened in the treatment period between 2016 and 2019 was on the pink line. *Violet*: Police jurisdictions where the first metro station that opened in the treatment period between 2016 and 2019 was on the violet line.

A.4 Robustness Checks

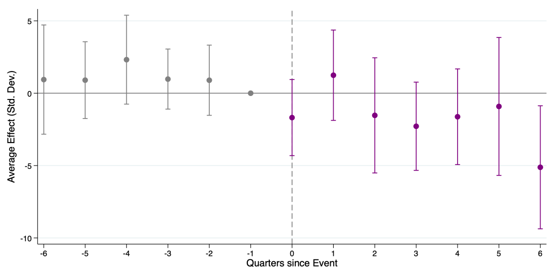
A.4.1 Alternate Comparison Group

Table A3: Comparison Group - Never Treated Jurisdictions

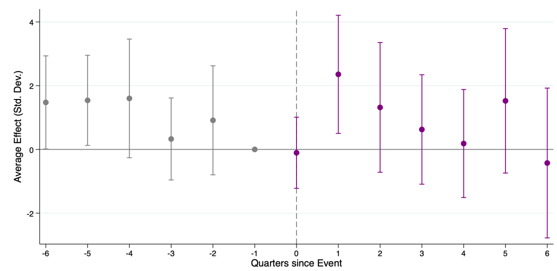
	HoS Crimes Against Women	Non-Gender-Specific Street Crimes	All Crimes Against Women	Domestic Violence
Post 1^{st} Metro Station	$\mathbf{-1.963}^{*}$	4.293	-2.174	0.101
	(1.094)	(5.085)	(1.366)	(0.630)
Observations	1,264	1,264	1,264	1,264
Jurisdiction Fixed Effects	Yes	Yes	Yes	Yes
Quarter Fixed Effects	Yes	Yes	Yes	Yes
P-value	0.073	0.399	0.112	0.873

Notes: This table presents the dynamic ATT following Callaway & Sant'Anna (2021) of the opening of the first metro station in a police jurisdiction. The comparison group includes never-treated jurisdictions. The time frame consists of six quarters before and after the treatment. Standard errors are clustered at the police jurisdiction level. Fixed effects are for each police jurisdiction and each quarter between 2016 and 2019. ***, **, and * indicate significance at the 1, 5, and 10 percent critical levels.

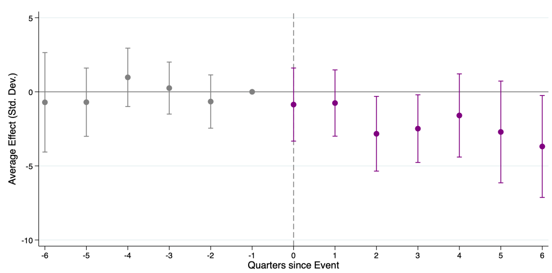
Figure A10: Even Study Plots: Comparison Group - Never Treated Jurisdictions



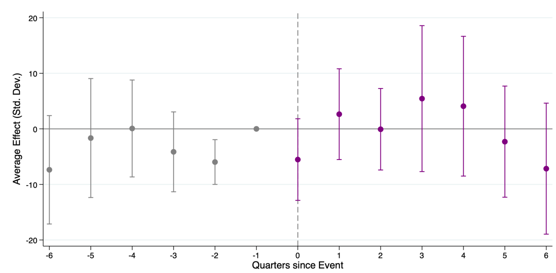
(a) All Crimes Against Women



(b) Domestic Violence



(c) **HoS Crimes Against Women**



(d) Non-Gender-Specific Street Crimes

Notes: This figure presents the event study coefficients and the 95% confidence interval following Callaway & Sant'Anna (2021) of the opening of the first metro station in a police jurisdiction. The comparison group includes never-treated jurisdictions. The time frame consists of six quarters before and after the treatment. Standard errors are clustered at the police jurisdiction level. Fixed effects are for each police jurisdiction and each quarter between 2016 and 2019.

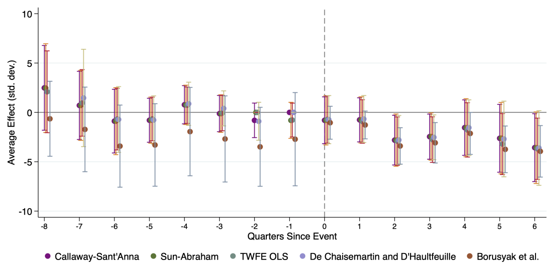
A.4.2 Other Estimation Methods

Table A4: Different Estimation Methods: Comparison Group - Never Treated Jurisdictions

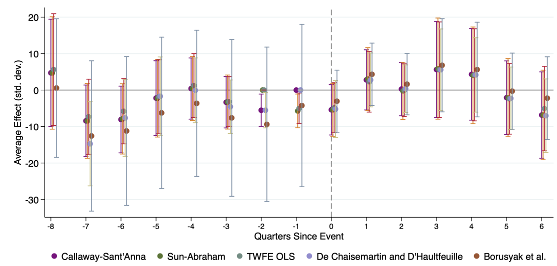
	HoS Crimes Against Women	Non-Gender-Specific Street Crimes	All Crimes Against Women	Domestic Violence
Post 1^{st} Metro Station				
Callaway & Sant'Anna (2021)	-1.963^{*}	4.293	-2.174	0.101
	(1.094)	(5.085)	(1.366)	(0.630)
Sun & Abraham (2021)	-2.094^{*}	0.672	-1.695	0.689
	(1.207)	(4.414)	(1.491)	(0.646)
de Chaisemartin & D'Haultfœuille (2020)	-2.019^{**}	-0.437	-1.547	0.789
	(1.029)	(4.017)	(1.382)	(0.691)
Dynamic TWFE	-2.323^{*}	0.540	-2.054	0.594
	(1.198)	(4.323)	(1.411)	(0.581)
Observations	1,264	1,264	1,264	1,264
Jurisdiction Fixed Effects	Yes	Yes	Yes	Yes
Quarter Fixed Effects	Yes	Yes	Yes	Yes

Notes: This table presents the dynamic ATT following Callaway & Sant'Anna (2021); Sun & Abraham (2021); de Chaisemartin & D'Haultfœuille (2020) and dynamic TWFE of the opening of the first metro station in a police jurisdiction. The comparison group includes never-treated jurisdictions. The time frame consists of six quarters before and after the treatment. Standard errors are clustered at the police jurisdiction level. Fixed effects are for each police jurisdiction and each quarter between 2016 and 2019. ***, **, and * indicate significance at the 1, 5, and 10 percent critical levels. Estimates from Borusyak et al. (2021) are not reported because they may be unreliable, and the standard errors may be downward biased due to insufficient sample size.

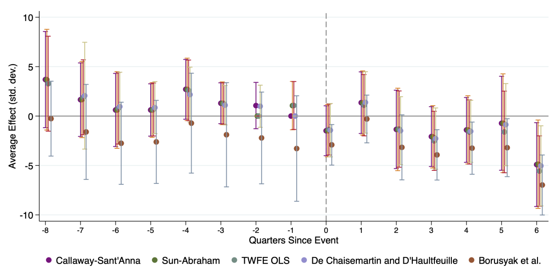
Figure A11: Even Study Plots: Different Estimation Methods



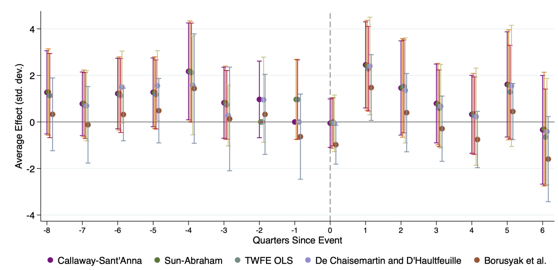
(a) **HoS Crimes Against Women**



(b) Non-Gender-Specific Street Crimes



(c) All Crimes Against Women

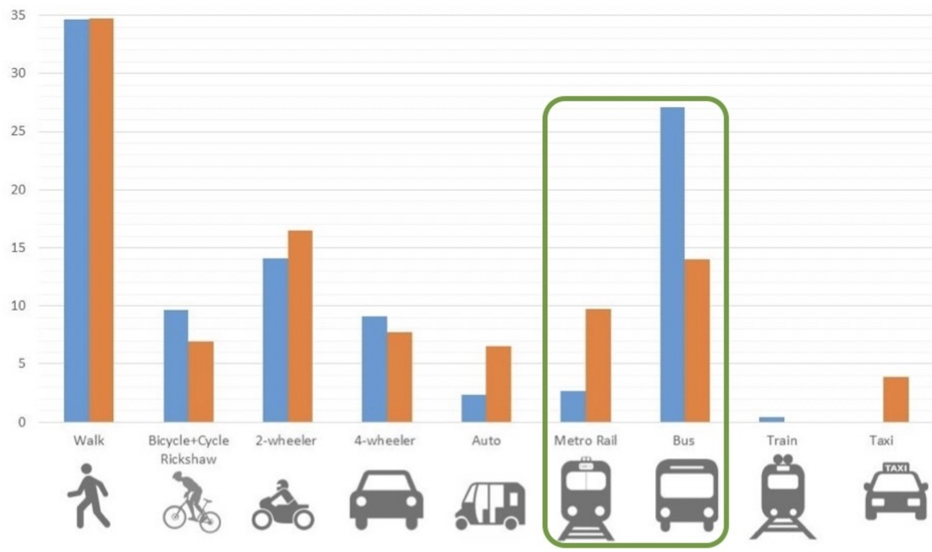


(d) Domestic Violence

Notes: This figure presents the event study coefficients and the 95% confidence interval following Callaway & Sant'Anna (2021); Sun & Abraham (2021); Borusyak et al. (2021); de Chaisemartin & D'Haultfœuille (2020) and dynamic TWFE of the opening of the first metro station in a police jurisdiction. The comparison group includes never-treated jurisdictions. The time frame consists of eight quarters before and six quarters after the treatment. Standard errors are clustered at the police jurisdiction level. Fixed effects are for each police jurisdiction and each quarter between 2016 and 2019.

A.5 Safety Effect

Figure A12: Modal Split for Different Transportation Alternatives (2007 and 2018)



Notes: Source: National Institute of Urban Affairs, 2020.