Street Police Patrols and Gender-Based Violence in Public Spaces: Experimental Evidence from Urban India

Pre-Analysis Plan

Sofia Amaral[1], Girija Borker[2], Nathan Fiala[3], Anjani Kumar[4], Nishith Prakash[5], Maria Micaela Sviatschi[6]

Abstract: How can gender-based violence in the public sphere (GBV) be prevented? Can improved police presence help curb street harassment? What works in improving victim's engagement with police services? This project aims to answer these questions through a novel policing program in Hyderabad, India. The Safety, Health, and Environment (SHE Teams) Program is a hotspots street police patrolling intervention targeting GBV in public spaces. The researchers and Hyderabad City Police have jointly developed research that aims to test the role of increased police presence through patrolling and policing visibility (i.e., uniformed vs. undercover officers). Our research will address fundamental questions in the economics of crime and gender.

Keywords: Hotspots policing, gender-based violence in public spaces, physical mobility, gender

---

[1] Ifo Institute, Ludwig-Maximilian University of Munich and CESifo, amaral@ifo.de.
[2] World Bank, gborker@worldbank.org.
[3] University of Connecticut and RWI Essen, nathan.fiala@uconn.edu.
[4] Hyderabad City Police.
[5] University of Connecticut, nishith.prakash@uconn.edu
[6] Princeton University, msviatschi@princeton.edu

1. Introduction

In India, gender-based violence (GBV) is a major epidemic and a significant public policy problem since it negatively impinges on women's physical mobility, education choices, and labor force participation (Borker, 2018; Siddique, 2018). A recent cross-country survey reveals a 50% lifetime victimization of sexual harassment in public spaces (Livingston, 2015).

Despite being such a fundamental public policy concern, interventions that address GBV in public spaces are limited (Moser, 2012). Further, the police are one of the main forms of government representation. They are essential stakeholders in addressing GBV in public spaces yet, little is known about its effectiveness at detecting, punishing, and preventing the problem. Recent research suggests that improving access to law and enforcement for women can reduce rates of domestic violence and improve standards of reporting GBV crimes.

This pre-analysis plan is to layout our strategy to study a project where we aim to understand the causal impact of improved policing for sexual harassment in public spaces. To do so, we have partnered with the Hyderabad City Police (HCP) to evaluate their flagship program to address GBV in public spaces, The Safety, Health, and Environment (SHE Teams) Program. The SHE Teams operates through a GBV-targeted street policing patrol to detect GBV in public spaces and engaging with citizens.

This study seeks to answer the following two main research questions:

1. What is the effect of GBV-targeted street patrolling on the frequency and type of street harassment incidents?
2. What drives these changes? Visible state presence and quantity of focused task-police force?

To address these research questions, we will make use of a clustered experiment to evaluate the efficacy of different forms of policing in curbing GBV crimes. First, we evaluate the effect of GBV-targeted street patrolling on the type and frequency of street sexual harassment (SH) and women's proactive response through increased police presence. Second, we evaluate the effects of police visibility in reducing the incidence of GBV crimes. The intervention involves two treatment arms – Visible, patrolling by SHE Teams in uniform and Undercover, patrolling by SHE Teams in civvies. The research design will allow us to disentangle the direct deterrence effects generated by police presence separately from an additional impact of visible specialized police presence. In the context of GBV per se, the "police visibility" component has not been studied, and it becomes even more relevant given the low access women have to police services. Lastly, we also seek to evaluate what forms of policing shift women's beliefs and choices regarding access to police. There is limited evidence on how improving policing towards GBV crimes may impact citizens' engagement with the police,

victimization, and detection of such crimes - key accountability outcomes for the police. Understanding such issues are a crucial steppingstone to addressing GBV, improving the accountability of the police, and changing citizen's engagement with one of the essential components of governance, and this is the focus of our research project.

2. Description of Treatments and Experimental Design

We use a clustered experiment to identify the effects of different approaches to policing. The unit of intervention for this study is a hotspot. Hotspots are public places like schools, colleges, bus stops or marketplaces where incidence of sexual harassment is relatively high. The research team in collaboration with the HCP identified 350 hotspots across Hyderabad which were randomly distributed across three groups - see Table 1 for a summary. The first group – Treatment arm 1 or the Undercover arm consists of 100 hotspots, which receive undercover patrolling, i.e., patrolling by SHE Teams in civilian clothes. The second group – Treatment arm 2 or the Uniform arm consists of 100 hotspots, which receive patrolling by SHE Teams in their police uniform. The third group is the control arm which consists of 150 hotspots which receive no patrolling through the course of the intervention i.e. receive the typical police patrol any other area would receive. Under the two treatment arms, each SHE Team is comprised of three officers, with at least one female officer.

To answer the first research question, does GBV targeted street patrolling effect the type and frequency of street sexual harassment, we will compare the control group to the undercover treatment group. For the second research question, is the effect on GBV crime driven by visibility of police, we will compare the uniform arm to the undercover arm. To understand whether victim's initial engagement and perpetrators behavior are different as a function of police exposure, we will use the fact that some police patrols will be undercover and therefore undetected by victims wishing to seek protection or advice. Conversely, teams of uniformed officers are more visible and therefore can signal to victims and perpetrators that police presence is higher in each location. In this way we will be able to isolate the detection effects from reporting effects.

Table 1: Summary of the Design

	Treatment Type		
	Arm 1	Arm 2	Control group
Team Type	Uniform Teams	Undercover Teams	"Business as Usual"
Number of Hotspots	100	100	150

2.1 Randomization

A stratified clustered randomization was done for the 350 hotspots across the city of Hyderabad. The stratification was done on the following basis:

- Nature of hotspot: Hotspots were categorized into four types – Educational (schools and colleges), Commute (bus stops and railway stations), General public places (markets and temples) and Residential – as the hotspots would differ in their composition of population (age and gender) basis their nature; and
- Population density: Hotspots were categorized into four categories – low, medium, large and very large – on the basis of the footfall at the hotspot. This stratification parameter was used because frequency of sexual harassment may be affected by how densely or sparsely occupied a location is. A footfall of less than 30 was termed as low; a footfall between 30 and 150 was termed as medium; between 150 and 400 as large and anything above 400 as very large.

A total of 2,000 iterations were done on a total of 57 key variables using data collected at baseline to arrive at a successful randomization. The three groups were balanced across observable characteristics of women like age, education, occupation, marital status, general safety perception and safety perceptions at various parts of the day. The randomization was implemented on Stata.

3. Data

This project will use three main data sources. We describe each of these data below.

3.1. Observational Data

We collect data on the prevalence of sexual harassment using an observational data collection exercise. To do so, we recruit female enumerators who observe various forms of harassment faced by other women using a mobile application at different times of the day at hotspots. Each enumerator covers 6 locations in a day spending 15 minutes at a location and then commutes from one location to another using public transport. During this time they record all instances of GBV that they observe, identify the type of harassment and the action taken by the victim or bystanders, if any. This will allow us to have an ongoing measure of harassment experienced by women at different times of the day in Hyderabad for 6 months independent of reporting effects and multiple records per each enumerator.

3.2. Survey Data

To measure proactive behaviors regarding GBV and victimization rates of GBV in public spaces, we will make use of a baseline and an endline victimization survey. The baseline survey was conducted with an estimated sample size of 10 women per hotspot. To evaluate if increased patrolling at hotspots leads to the displacement of GBV crime, we identified two neighboring areas or spillover areas per hotspot with the help of the HCP – one within a radius of 200 meters and another within 500 meters. Roughly, five women were surveyed per spillover area. The baseline data collection was done on a sample of 8,264 women.

This survey will be followed up by an endline victimization survey which will be conducted at the end of the intervention. The survey will include questions related to own and peer experiences of public space violence. The estimated sample size at endline is 13,000 women, depending on budget requirements.

3.3. Administrative Data

We will use information on crime reports through emergency calls, police walk-ins and red-handed cases. This information is at the report-level, it is geolocated and with the precise date of the report. We will also use HCP's dense network of CCTV cameras to obtain a unique measure of female mobility. We will make use of machine learning and artificial intelligence techniques to generate the count of females visible at the hotspots. This information will be used at the hotspot - or spillover unit level.

4. Key Variables and Measurement

4.1. Primary Outcomes

The primary question that this study seeks to answer is the effect of street patrolling on the incidence of sexual harassment and women's proactive responses. In line with this, we will use the following primary outcomes:

- **Incidence of street sexual harassment:** We capture data on the incidence of street sexual harassment through an Enumerator Observation Survey. We ask enumerators to observe a hotspot for 15 to 20 minutes every day for a duration of six months and record all instances of street sexual harassment observed. Given the social stigma attached with victimization of sexual harassment, we may expect women to underreport their experience of street sexual harassment. Therefore, to circumvent this

problem, we will use the 'observed' measure of the incidence of sexual harassment as captured by the Enumerator Observation Survey (EOS). The EOS provides us with a hotspot-level incidence of sexual harassment.

- **Safety Perception:** Lastly, we seek to study the effect of street patrolling on safety perception of women. We will obtain this measure from the victimization survey. Women are asked the following questions to arrive at a comprehensive measure of safety perception:
    - How safe they feel at the hotspot?
    - How safe they feel about traveling anywhere in the city between (i) 6 AM and 12 PM; (ii) 12 and 4 PM; (iii) 4 and 10 PM

Responses are collected on a four-point scale: 'Completely Safe', 'Somewhat Safe', 'Somewhat Unsafe' and 'Completely Unsafe'. Through this variable, we will address whether women report feeling safer in areas patrolled by the police.

- **Mobility:** As a part of women's proactive responses, we would like to see if increased police presence leads to increased mobility of women. Here, we define mobility as the footfall of women at a hotspot at any given point of time during the day. To measure the footfall of women, we will leverage the dense network of CCTV camera of the Hyderabad City Police. Making use of artificial intelligence and machine learning, we would arrive at a measure of footfall of women at the hotspot as captured by the CCTV cameras.

- **Routes taken to work:** In the victimization survey, women are asked about the routes they take to reach their workplace/colleges from their homes. Safety concerns may lead women to take inefficient routes (in terms of time/ distance and cost) to reach their workplace or colleges. Through this data, we can analyze whether women alter their routes to workplace/colleges in response to increased patrolling.

4.2. Secondary and Exploratory Outcomes

Apart from the type and frequency of street sexual harassment, street patrolling may also affect other aspects of women like reporting behavior, participation in social activities or travel routes taken by altering their safety perception about locations. Therefore, we will look at the following variables.

- **Change in type of harassment:** We hypothesize that increased street patrolling will lead to a decrease in the incidence of street sexual harassment. However, we would also like to see if increased patrolling has differential effects on different forms of

harassment experienced by women. Data from the EOS will be used to analyze the change in the type of harassment observed at the hotspot, if any.

- **Reporting of incidence to police or other changes in pro-active behaviors:** Findings from the baseline victimization survey indicate that the reporting rate of street sexual harassment is only 7.5 percent. Increased police patrolling, especially in the visible arm will increase the visibility of police. We thus hypothesize that reporting rates of GBV-related crimes may increase with increased police visibility as women may find it easier to approach police. We intend to test this hypothesis using data on reporting rates from the victimization survey. To measure other proactive behaviors, we will construct - through the observational data - a measure as to whether victimized women react more proactively. This will be done through question answers (1/0) to whether "Does the victim..." did any of the following: i) "called someone over the phone to help", ii) "called out the perpetrator directly", , iii) "informed the person accompanying her at the location", iv) "use any form of self-defense", or v) "ask for help from bystanders".

- **Incidence of harassment while traveling:** As a part of the EOS, enumerators also make observations on sexual harassment while traveling between two hotspots, i.e., in public transport like buses, shared autos, trains, etc. This enables us to analyze whether the effect of hotspot-level police patrolling on incidence of sexual harassment is local, i.e., it reduces sexual harassment only at the hotspot-level, or whether it reduces sexual harassment faced by women in public transport as well.

- **Labor Market Choices:** To measure if increased police presence leads women to modify their labor market choices, we will analyze the 'hours' spent at workplace. This variable will be obtained from the victimization survey where women are asked to report their time of arrival/departure to/from the workplace.

- **Participation in social activities:** An improvement in safety perception of public places may induce women to step out more frequently and may also alter the time at which they are present in public places. For example, street police patrolling after in late evening may enable women to step out in late evening because a sense of safety induced by police presence. To record this, as a part of the Victimization survey we ask women whether they engage in any of the following social activities:
  – Going out for a movie with female friends – Shopping for rations
  – Attend recreational classes
  – Visit relative's home
  – Got out to explore the city

Further, we also ask them about the time at which they carry out each of these activities.

Through the baseline and endline victimization survey, we will capture the change in the participation in activities as well as the time at which these are conducted.

4.3. Other Outcomes

In addition to the effect of street patrolling on victims of sexual harassment, we would also like to study the effect of street patrolling on the perpetrators.

- **Number of people caught by the police: (Red-handed cases):** As a part of their patrolling program, SHE teams catch criminals red-handed and a petty case is filed against the perpetrator. We would look at the change in the number of red-handed cases over the course of the intervention to assess the effect of increased patrolling on the number of red-handed cases.

- **Reported cases to the police:** To corroborate the self-reported measure of reporting of street sexual harassment to police, we will use the data from police on complaints registered to see the change in reported cases to police. For this purpose, we will use complaints registered using Dial 100, a police emergency service hotline number used for reporting crimes.

- **Compliance with the randomization and treatment:** A large part of our project requires a good compliance by the partners involved - the police. To understand this, we plan to test for compliance with the randomization and treatment through GPS tracking of patrolling vehicles.

5. Empirical Strategy

This section describes the empirical strategy we will be used to test the hypotheses generated above.

5.1. Intent to treat Effect

The effects will be estimated through an intent-to-treat estimator (ITT). We will follow an empirical strategy whereby we will estimate program impacts on the outcomes outlined in Section 4 by calculating the ITT via OLS. For the primary outcome using the observational data we will estimate the following equation:

$$Y_{wh} = \alpha + \beta T_{wh} + \gamma X_h + \delta_{wh} + \epsilon_{wh} \qquad (1)$$

where Y denotes the outcome for a harassment at the hotspot during week w; T is an indicator for assignment to treatment and β is the coefficient of interest. We will also include a vector of baseline covariates and a hotspot fixed effects δ_h, required because the probability of

assignment to treatment varies by strata); and ϵ_h is a hotspot error term clustered by hotspot. For the survey data, we will estimate:

$$Y_{ih} = \alpha + \beta T_h + \gamma X_{ih} + \delta_h + \epsilon_{ih} \tag{2}$$

where Y denotes the outcome for person i in hotspot h; ITT is an indicator for hotspot assignment to treatment; X is the set of baseline covariates for women i; a vector of baseline covariates for women i and hotspot h and a hotspot fixed effect δ_h. The error term ϵh is a person i hotspot h error term clustered by hotspot. We will be exploiting: i) the panel structure of the data collection with women at baseline and endline, and ii) exploiting the cross-sectional variation in treatment status in the women's survey and the endline data collection. As per the choice of control variables we will use the post-double selection lasso to choose the controls. We also plan to examine heterogeneous treatment effects by conducting heterogeneity analysis across two key parameters we predict to be on research interest pertain to: i) the differential role of population density by hotspot; ii) norms index.

5.2. Correction for Multiple Hypotheses

To correct for multiple hypotheses testing, we will follow Katz, Kling and Liebman (2007) in calculating bootstrapped estimates of adjusted p-values using a modification of the free step-down algorithm of Westfall and Young (1993). We will correct for multiple hypothesis tests in the primary outcomes only.

6. Timeline

The baseline data collection took place in August 2019 and the intervention was launched on 16th September 2019, after achieving a successful randomization. The intervention is expected to run for a duration of six month up to March 2020. Endline data collection is scheduled to take place between March and April 2020.

The police have decided to take lessons they have learned from the initial implementation of this program and change their operations slightly, starting in December 2019. The police will now include posters announcing the presence of the SHE Team in undercover locations. They will also move their policing times in some hotspots slightly so as to better target when GBV is taking place. Data from the EOS will allow us to observe if there are immediate effects from this change, but as there will be no experimental variation, we will not be able to make strong causal statements about the change. The endline victimization survey will only cover effects from the change in deployment strategy.

7. Partnership and Human Subjects

The project has received the approval from the IFMR ethics committee, JPAL ethics committee and Princeton University ethics committee. The project has also received approval from the Commissioner of Police, Hyderabad City Police to conduct the survey and receive required support from the police.