Impact Evaluation Concept Note Template

The Costs of Sexual Harassment in Public Transit in Rio de Janeiro, Brazil

Brazil

May 2017

This is a trimmed-down version of the original Concept Note approved for the project, which with classified as official-use only. This version contains only selected sections that are relevant as a record of the study's analysis plan.

Table of Contents

1. Executive Summary

Sexual harassment imposes psychological and economic costs: it may affect women's decisions on how and when to move in the public space, potentially affecting significant decisions such as job choice. However, quantifying the impact of harassment is difficult because of measurement issues. This study uses a novel experiment to quantify the economic impact of harassment in urban public transport, a setting where harassment is pervasive. We recruit women who commute on Rio de Janeiro's Supervia, a suburban train system, to report on their rides on a smartphone app in return for compensation. We randomly assign them to ride on a train carriage designated for women only to test whether this reduces their experience of harassment. In a separate stage of the experiment, we offer these women the option to choose which carriage to ride, varying the payment across types of cars to estimate their willingness to pay to ride the women-only car. We gather evidence to test whether willingness to pay is due to harassment or other explanations such as different levels of crowding or reduced exposure to general crime. Unlike most other work in this area, our study will help to quantify the costs of the specific threat of sexual harassment in the public space, which is widespread and frequent.

2. Background and Key Institutional Features

Harassment on public transport is shockingly common in countries across the developing and developed world. This is also the case in many major cities of Brazil. A recent survey in São Paulo suggests that public transport is the most common place where women suffer harassment: 35% of female respondents reported being harassed either physically or verbally while using public transport (Datafolha 2015). Underreporting likely makes these statistics an underestimate of the problem. Access to safe public transport is a critical barrier for women's access to economic opportunities. Given their multiple responsibilities as well as negative experiences on public transport, women often choose jobs on the basis of distance and ease of travel, choices that tie them to local work options. These limitations are particularly severe for poor women, who often reside in more marginal neighborhoods where most available jobs are informal and low in productivity. Investments in safe public transportation can increase women's access to economic opportunities by reducing travel time and increasing mobility.

Our study will focus on Rio de Janeiro, where this pattern also holds: most low-income families reside in the periphery of the Rio de Janeiro metropolitan region (RJMR), while jobs are mainly concentrated in the Municipality of Rio de Janeiro. The RJMR has an extensive public transport system including a metro, suburban rail and bus-rapid-transit system (BRT). A user survey conducted by the Rio de Janeiro BRT suggests that security is indeed among the main concerns for commuters. As part of the design of the study we held focus group discussions with women commuters, which suggested women endure unwanted physical contact on average once a week.

Our study focuses on Rio de Janeiro's main suburban rail system, the Supervia. This system connects downtown Rio with its outskirts, including many low-income areas. The system has seven lines which all terminate within at the Central do Brasil station "Central" in the center of Rio de Janeiro (Figure 1); most users travel from the outskirts to Central in the morning, often then changing to another mode such as

the metro or bus, and return in the opposite direction in the evening. The Supervia carries around 700,000 passengers a day; on average half the passengers on the system are women.

Figure 1: Supervia map

This study exploits a law passed in 2006 which obliged both the Metro and Suburban train operators to offer women-only cars during rush hours (6-9AM and 5-8PM), when crowding is heaviest. The law does not specify the amount of space that needs to be reserved. Supervia reserves one car for women in each train composition, which in most cases means 1 of 8 cars. At other times of the day these cars serve as mixed-gender cars. Hence, our study focuses on these rush-hour time periods.

The women's-only rule is enforced by platform guards who also have the general responsibility of managing the boarding process. The presence of these guards varies substantially by station, and is particularly low in stations in the outskirts of the metropolitan area. The guards do not have policing power, further hampering effective enforcement of the law. As a result, male compliance with the policy varies substantially geographically. In addition, the carriages are connected internally, so it is possible for men to move from other cars to the women's-only car after boarding. Estimating the value women assign to riding women-reserved cars is all the more relevant in this context that the municipal government

intends to expand enforcement of the women's-only rule on the system by 1) increasing fines for men who do not comply with the original law and 2) including police forces as part of the enforcement body.

Figure 2: Women-only car

Figure 3: Non-compliance with women-only car rule

3. Literature Review

There is a broad literature on the economic consequences of crime and violence, with a subset focusing on the effects on women or addressing gender-based violence. Most estimates of the costs of crime or gender-based violence have typically taken an accounting approach (cf. McColister et al. 2010, Duvvury et al. 2013 for reviews). This approach adds up estimated costs such as health care and missed work, but fails to capture indirect effects such as fear of violence or steps taken to avoid violence. This is better addressed by studies examining the consequences of insecurity on economic activity, such as Velasquez (2014), who uses panel data from Mexico to examine the effect of waves of drug gang related violence (which often targeted women). She finds that when such violence increased in an area, women withdrew from the labor market. Several studies look specifically at violence in the public space on mobility and resulting economic consequences. Jacoby and Mansuri (2011) use an instrumental variable strategy based on school placement in high-population settlements to demonstrate that travel outside of the neighborhood of one's caste settlement reduces the likelihood of girls attending school in rural Pakistan, but does not have a similar effect for boys. They argue and show with qualitative evidence that this is due to lack of safety for girls when leaving the area dominated by one's own social group. Muralidharan and Prakash (2013) show that the impact of providing cycles to girls in Bihar, India on schooling is significantly larger than that of similar conditional cash transfers, suggesting that improved safety from girls cycling in groups is the mechanism for the large effect of the program. Cheema et al. (WP) find that transport provided by a trusted member of the community for groups of women is far more effective at increasing women's participation in vocational training than an equivalent cash transfer to cover the cost of transportation.

As Jayachandran (2015) highlights, constraints on women's physical mobility in the public space are widespread, but it is difficult to distinguish to what extent these constraints are motivated by concerns about safety versus controlling their "purity". Little work focuses on the specific mechanism of harassment, in part due to the challenges of measurement. Much of the work which does focus on specific types of violence looks at the most severe crimes, such as assault, and on the effects of women pulling out of economic activity (labor force participation or schooling) altogether. Our study will help to quantify the costs of the specific threat of sexual harassment in the public space, which is a lower level crime but more widespread and frequent. We are able to capture the costs which are incurred by women who do remain in the public space or economically active. The most closely related work to this project is that of Aguilar and Gutierrez (WP), who use a regression discontinuity design based on the timings of the women's-only car in Mexico City's metro system. They find that the imposition of the women's car decreases sexual violence against women while increasing pushing and shoving amongst men, possibly due to increased crowding. They use hypothetical questions to quantify willingness-to-pay for a completely safe ride, finding that the implicit cost of sexual harassment is substantially higher than that of pushing and shoving. Our study builds on their work by capturing more precise measures of welfare, including (a) experiences of physical and verbal harassment, (b) revealed preference rather than stated preference measure of willingness to pay. In addition, our design allows us to explore mechanisms, including (a) congestion, (b) compliance with the rule (c) variations in perception of risk, and (d) norms of women's versus men's responsibility for avoiding harassment.

4. Policy Relevance

The results of this study can shed light on the effects of the women's-only car policy and the government's efforts to increase compliance with this policy. Governments in a diverse set of contexts have implemented women's-only spaces on public transport vehicles. Cities in Brazil, Mexico, Pakistan, India, Bangladesh, Iran, Egypt, the UAE, Israel, Philippines, Malaysia, Indonesia, South Korea, and Japan have such policies. However, these policies are highly controversial. Opponents argue that they reinforce norms of gender segregation, place responsibility for avoiding harassment with women rather than addressing the root problem, and increase stigma and harassment for women who choose to ride in the mixed-gender spaces. Furthermore they could create negative spillovers in the mixed spaces if women who chose not to use the women-only space are perceived as not minding the harassment. Supporters argue they allow women to travel without the threat of harassment, encouraging them to travel freely when they otherwise might not, or to travel on more sustainable and affordable modes of public transport when they otherwise would choose private transport. Ultimately, if women are more comfortable traveling they may be able to take up a wider range of employment opportunities, as well as accessing the public space more broadly. While these spaces do not address the root problem of harassment, supporters argue, they are an effective second-best solution that is feasible in the short term. To date there is no systematic evidence on the effects of gender-segregated spaces for women on public transport, and decisions on this controversial issue are made ad hoc based on the perceptions of the implementer. While our study cannot capture the possibility of displacement of harassment to the mixed-gender car, it will bring evidence on other aspects of the impact of the gender segregated car to a policy debate that is heated and yet poorly documented.

However, the study also serves a broader purpose: to use the women's-only car policy on Supervia as an experimental setting to quantify the economic and psychological costs of sexual harassment in the public space. While harassment is known to be widespread, particularly on public transport, to our knowledge the costs have never been quantified. The results will help to motivate piloting and evaluation of the range of policies and interventions, beyond gender segregation, that can be used to try to address this problem. These include:

- Training and incentivizing transport and security staff to enforce anti-harassment laws and rules;
- Interventions targeted at perpetrators, such as "name and shame" interventions;
- Interventions targeted at potential victims of harassment, designed to encourage and empower them to speak up and/or report harassment when it occurs;
- Interventions targeted at the general public, designed to raise awareness of the problem and encourage bystanders to intervene when they witness harassment.

Rigorous evaluation of interventions to address gender-based violence are limited to date (Kiplesund and Morton 2014). Both the experimental methods developed in this study and the partnership developed with Supervia can be used as a basis for future impact evaluation of such interventions in the Rio context.

5. Theory of Change

This study uses assignment to the women's-only car as an exogenous source of variation in the presence of men in the ride environment. This relationship is measured by a mapping team which measures the number of men on the women's car. The enforcement pilot with Supervia (phase 4), if successful, can also be used to exogenously vary this presence even in areas where baseline compliance with the women's-only car was low.

Decreased presence of men on the selected car may lead to decreased harassment on that car (not necessarily overall). This may happen through increased social enforcement (potential harassers are more visible and less likely to be anonymous in a crowd), through sorting (potential harassers go to the mixed car) and/or through simple numbers – without sorting there would simply be fewer men to harass women. We can observe reductions in harassment in Phase 3, when women are randomly invited to ride the women's or the mixed car, and report on their experiences including specific experiences of harassment. We cannot identify which mechanisms are operating.

If harassment is decreased, the reduced emotional and psychological burden on female riders (measured in Phase 3) may then lead to increased willingness to pay (measured in the revealed preference exercise in Phase 2). Each step in this causal chain will be measured in Phase 3, and their psychological wellbeing.

Willingness to pay represents an economic quantification of the reduced emotional and psychological burden from harassment. This is an economic cost of harassment in its own right. Harassment may also lead to lower willingness to travel more often or on times / routes that are optimal for the individual's commute (measured through adjustment on these margins in phase 3). Finally it may result in the choice to avoid travel completely or switch to a completely different mode of travel, which ultimately may affect outcomes such as job choice and schooling; this is beyond the scope of this study.

Alternatively, the reduced presence of men could affect WTP through two other possible mechanisms: (1) reduced crowding in the women's car, which is measured by the mapping team; and / or (2) reduced overall crime, which is measured as an outcome in Phase 3.

One potential unintended consequence of the women's-only car is that of spillovers due to the sorting mechanism described above: potential harassers move to the mixed car. This does not affect the main analysis of the costs of harassment, because that relies only on a change in the probability of harassment within the women's car to identify this cost. However, it does have important implications for the women's-only car policy itself. Measurement of these spillovers is beyond the scope of the main experimental components of the study, but we will attempt to capture them in focus groups and the platform survey.

A second potential unintended consequence is the reinforcement of the social norm that women are responsible for avoiding harassment. This could potentially act as an alternative mechanism for willingness to pay if women are willing to pay to avoid social stigma. We will capture the mechanism of harassment through Phase 3, and we will attempt to capture potential social stigma through representative platform survey data, including Implicit Association Testing on these issues.

6. Hypotheses/Evaluation Questions

The main question posed by this IE is: How much are women willing to pay to avoid harassment? The secondary evaluation question is: what are the effects of the women's-only car on harassment? We will test each of the following key hypotheses to answer the main question:

- The women's-only car reduces harassment experienced by women who choose to ride it.
- This reduced harassment improves their psychological wellbeing.
- Women are willing to pay for this reduction in experienced harassment.

In addition, we include supplementary research methods (the IAT, platform survey and focus groups) to help document the role of social norms that assigns responsibility to women for avoiding harassment.

7. Main Outcomes of Interest

Table 1. Main Outcomes of Interest

	Outcome Name	Definition	Measurement Level
1	Take-up of women-only car	Binary variable on type of car taken	Individual/ride
2	Willingness to pay	Take-up of women-only car at different levels of opportunity costs (revealed) Take-up of women-only at different levels of opportunity costs (stated)	Individual/ride Individual (exit survey)
3	Well-being	Scale 1-10 How much do you feel [happy / sad / frustrated / stressed]?	Individual/ride Individual/ride
4	Occurrences of harassment	Did you experience different types of harassment (unwanted comments, touching or robbery)?	Individual/ride (ride data)

Table 2. Interaction variables

1	Compliance	Percentage of men in the women-only car	Individual/ride, line (ride data) segment/time bin (mapping data)
2	Congestion	Categorical indicator on congestion Load factor (passengers/m2)	Individual/ride, line (ride data) segment/time bin (mapping data) Station/month/hour (admin)
3	Perceived risk	Were you worried you would experience certain types of harassment?	Individual/ride (ride data)

	How likely is it that a women experiences certain types of harassment?	Individual (exit survey)

8. Evaluation Design and Sampling Strategy

8.1 Scope

Harassment can impose psychic costs on women who travel on public transport. If the potential cost is sufficiently high, some women may adjust their travel plans by changing their route or time to mitigate the cost. This study is designed to capture both the direct costs and the evidence of adjustments to avoid costs. Finally, however, women for whom the potential cost of harassment is high may even choose to opt out of public transport, choose not to take a particular job, or even travel at all on public transit systems. These individuals do not appear in the transport system and are beyond the scope of this study. (Since the women's-only car is an established and well-known policy in our setting, we may be more likely to observe such individuals using the transport system and resorting to the women's car than if it were newly introduced.)

8.2 Study setup and participants

SAMPLE A: 250 PARTICIPANTS SUBMITTING DATA THROUGH SMARTPHONES

We deploy a novel data collection approach in which we recruit participants to submit data through an Android app. Participants are asked to perform tasks that outline the exact data to be collected, location and pay-out for completion. The platform allows us to vary the content and pay-off for the task. In contrast to performing a user-survey, we are able to 1) observe contributors' use of the women-only car for multiple rides, 2) vary the pay-off to perform certain tasks and 3) randomize the tasks women are asked to do.

Contributors are recruited through Facebook, flyers at the stations and referral incentives for existing contributors. Upon signing up, participants fill-in a demographic survey. For the study approximately 600 contributors are recruited. Most contributors are frequent Supervia users and only women are included in the main part of the study. During the recruitment there is no specific mention about the study objective being related to the women-only car or harassment. The survey is framed as a general study into the user experience on the Supervia. Because the main respondent pool is a select sample, we will analyze the characteristics of our population as compared to the larger user base of the system and a platform survey. Once a participant signs up to the experiment, she is offered to get paid to ride the Supervia train on any weekday between 6-9AM or 5-8PM and answer a set of questions before, during and after the ride. Contributors are paid by completed task and free to opt-out at any moment. Therefore, the contract with

data collection agency Premise is based on the required number of participants completing the entire sequence of tasks. We can use data from initial activities to examine the correlates of opt-outt.

Each participant can ride up to twice per day (once in the morning, once in the evening) every weekday and ride from any station and on any line she chooses. Each rider will be given a set pipeline of tasks they can complete. The sequence of the assigned tasks will have participants go through the different phases of the study in a fixed order. With each phase the nature of the task differs slightly to enable us to answer the different research questions. The details of each phase are described below, and instruments are included in Appendix 1

PRE-PARTICIPATION SURVEY - COMPLETE

At baseline users are surveyed to gather demographic data as well as their residential location and typical use of the Supervia system.

PHASE 1: MAP COMPLIANCE AND BASELINE BEHAVIOR - COMPLETE

To date no systematic data has been collected on compliance and use of the women-only car. Anecdotal evidence suggests that male compliance with the policy is low and that there are large geographical differences in compliance in the system. The first phase of the study aims to trace current take-up of the different types of cars and gender composition.

In this phase women are paid one base amount to ride the suburban rail (Supervia) and report about their ride. They may freely choose each time to participate and if they participate, whether to ride the mixed or women-only car. In addition to the riders, a group of dedicated observers, a so-called mapping team, collect data on compliance with the women's-only rule and congestion from the platform. This ensures we have data on the trade-offs faced by women when deciding on which car to embark, instead of only observing the situation in the car they selected. The team is tasked to observed compliance and congestion in both types of cars at set time bins and line segments. The mapping team will continue their same platform based task for the remainder of the phases.

This will allow us to gauge how compliance and congestion affects take-up of the women-only car at baseline and at each premium level, and to address the alternative mechanism of differential crowding between the women's and mixed cars.

PHASE 2: QUANTIFY WILLINGNESS TO PAY BY ASSIGNING DIFFERENT PAYOUTS - COMPLETE

To quantify willingness to pay, we have designed a revealed preference experiment. In this phase contributors are still free to choose which car they take, but are offered an additional premium if they choose to ride the mixed car. For each rider we test three levels of premiums. Each contributor will perform multiple rides at each of the premium levels before moving to the next phase. This allows us to estimate the willingness to pay for the women's-only car, as the amount that the user is willing to forego

in order to choose that option. In addition, we can analyze how WTP varies under different contexts faced, while keeping user characteristics constant.

PHASE 3: QUANTIFY WELFARE BENEFITS BY RANDOMIZING CAR ASSIGNMENT - COMPLETE

In the next phase, participants will be randomly assigned to receive offers to ride either the mixed or the women's-only car (at the same price), and we will ask them to report measures of their experience and comfort level. This allows us to estimate the effect of riding the women-only car on measures of subjective well-being. Each contributor takes total of 10 rides at this phase. The order in which contributors are offered the women's car and mixed car is randomized. Every individual ride is always completely voluntary so women can choose not to participate on any day as well as not to continue, including deciding not to participate in mixed car rides if they prefer (this is an additional margin of adjustment which we measure). After each ride, the participant reports characteristics of the ride including any harassment experienced. In case she experienced harassment, the app directs her to resources available in the Rio area.

We minimize non-compliance by requiring a photograph from within the car, so participants must be in the correct car for at least part of the ride. We also capture some part of the non-compliance rate by asking a non-incentivized question in the task: "did you move to a different car?"

Figure shows how each of the tasks are presented in the application.

Figure 5- Tasks per phase

POST-PARTICIPATION SURVEY - COMPLETE

At the end of the study, users are surveyed to gather hypothetical willingness to pay questions which can be compared to the revealed preference results. In addition, they indicate their perceptions of risk of different kinds of harassment and crime on the mixed and women's car. The perceived risk of robbery is used to address the possibility that avoiding general crime is a potential alternative channel for WTP. Respondents are also asked for their views on the advantages and disadvantages of the women's-only car, which is further qualitative evidence on the mechanism associated with WTP.

SAMPLE B: REPRESENTATIVE SAMPLE OF 300 RESPONDENTS

REPRESENTATIVE PLATFORM SURVEY - PLANNED

To benchmark the results from our Premise sample with a larger random sample of Supervia user we conduct a complementary data collection of people recruited at the train platform. The survey will be short version of the exit survey and includes questions on women-only car ridership, stated willingness to pay, risk perceptions and perceptions of the policy. We use a simple sampling protocol based on ordering and counting individuals by platform location to ensure selection of a random sample representative of evening rush hour riders starting from the major hub station, Central do Brasil. Details of the sampling protocol as well as all draft instruments are included in Appendix 2. Pilots of the instrument have yielded a refusal rate of 6.7% for the survey, reducing the risk of again introducing a selection bias. We will be able to compare demographics from the platform survey participants with the characteristics of our Premise sample and administrative data from Supervia to gauge the representativeness of the sample.

IMPLICIT ASSOCIATION TEST - PLANNED

It is possible that the introduction of the women-only car has a negative spill-over effect by being perceived as a signaling device. For instance, women riding the mixed car after the introduction of the women-only care may be perceived as more open to being approached by men, after all they have the option of riding with only women. This could increase the risk of harassment for women who chose to take the mixed car. Some women may choose to ride the women-only car only to avoid sending any signal that they are "open" to harassment. If this were the case this would mean that women value the women-only car simply because now there is a potential stigma on women who ride the mixed car.

To explore this channel, we will conduct an Implicit Association Test (IAT) with respondents to the platform survey. The IAT will be adapted for this context: participants will be asked to classify words, stimuli, to categories that suggest women either being open or not to being approached by men, while being presented with pictures of women who are in the women-only or mixed car. The test will give us a sense whether men and/or women associate women with openness to being approached, depending which car they are in.

Each participant who consents to the IAT will participate in the IAT in a station on the platform, and will participate in a series of three IAT instruments in a randomly assigned order. We will calculate the IAT scores as detailed in Greenwald, Nosek and Banaji (2003).

- A ``safety'' IAT: A positive score means that the respondent associates users of the women's-only car in the Supervia system with seeking safety and users of the mixed gender car with less concern for safety.
- An ``openness to advances'' IAT: a positive score means that the respondent associates users of the women's-only car with being less open to advances from men, and users of the mixed gender car with being more open to advances.
- A ``gender-career'' IAT: a positive score means that the respondent associates women more strongly with home and men with career. This is a widely used tool which will be used as a benchmark.

We will test for differences in the IAT D-score between the safety and openness to advances IAT overall, and between men and women.

In piloting, approximately 40% of those interviewed in the survey agreed to participate in the IAT. In scaleup we intend to increase the payment for respondents from 20 reais to 40 reais to improve the response rate further. In addition, we will test whether responders to the IAT differ on observable characteristics from non-responders, including on survey questions covering similar content to that captured in the IAT.

ENFORCEMENT PILOT – PLANNED PENDING POLICY IMPLEMENTATION

The government has recently passed new legislation to improve compliance with the women's-only rule. With this change in policy on the table a potential second phase of the study is currently under discussion which will aim to evaluate the impact of the roll-out of the increased enforcement on compliance and welfare measures. The government intends to expand enforcement of the women's-only rule on the system by 1) increasing fines for men who do not comply with the original law and 2) making the operators responsible for enforcing the rules instead of merely offering a car that is meant to be women's only. However, the exact details on the exact implications in terms of enforcement mandate and timeline are still under discussion.

This study enables us to study new interventions in the system. We are working closely together with the train operator on roll-out of the enforcement to allow us to measure the impact of such an effort through differences-in-differences or an event study.

8.3 Sample Size Calculations

In advance of the Sample A data collection, we conducted power calculations by simulation to identify the tradeoff between the number of participants recruited and number of rides per participant. Figure 6 show results from a subset of these simulations, illustrating that we have approximately 70% power to detect an average effect size of 10-cents willingness to pay with 100 participants and at least 20 observations per participant, assuming between-participant (alpha) and within-participant (epsilon) random variation is distributed N(0, 0.5) – based on pilot data collection. Figure 7 shows similar results for an effect size of 15-cents willingness to pay, with power of 80% with 100 participants and 20 rides per participant at this effect size.

Individuals have the choice of discontinuing participation at any time. Therefore, the contract with data collection agency Premise was based on the required number of participants completing the entire sequence of tasks, and non-response would already be incorporated into these calculations.

Figure 6 – Power calculations for Sample A; effect size WTP = 10 cents

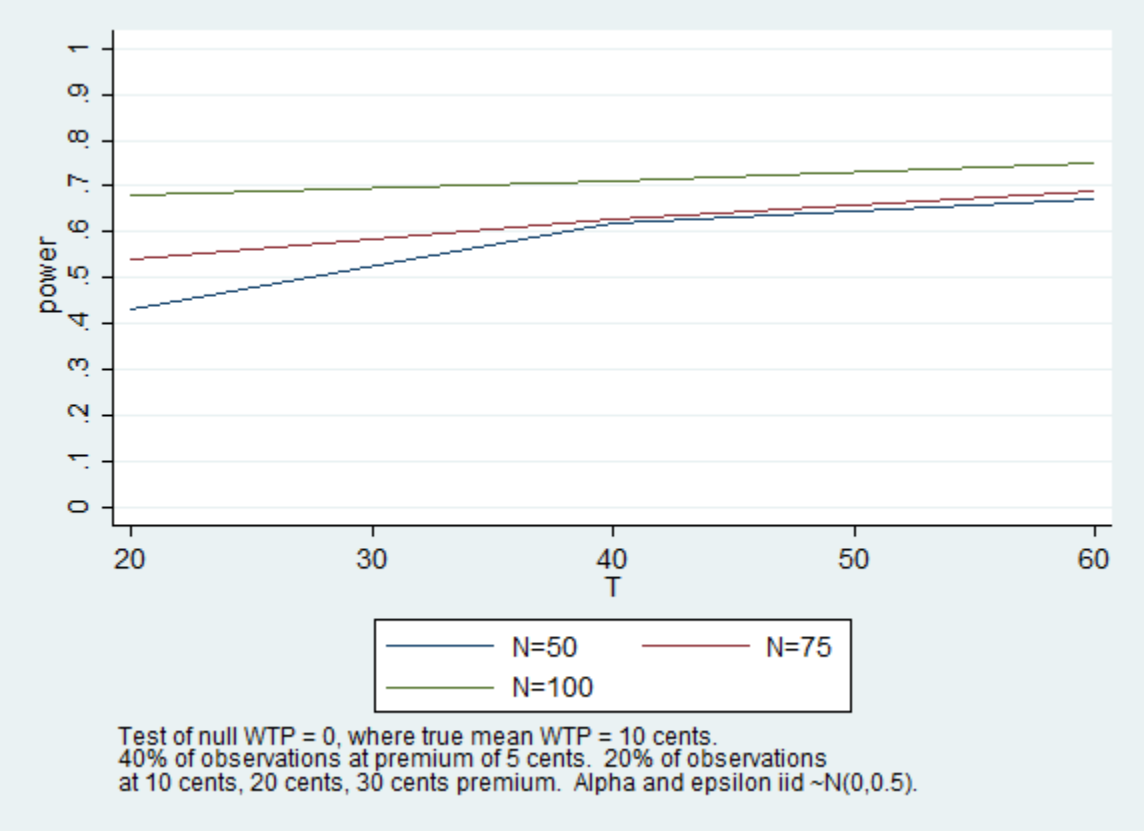


Figure 7 – Power calculations for Sample A; effect size WTP = 15 cents

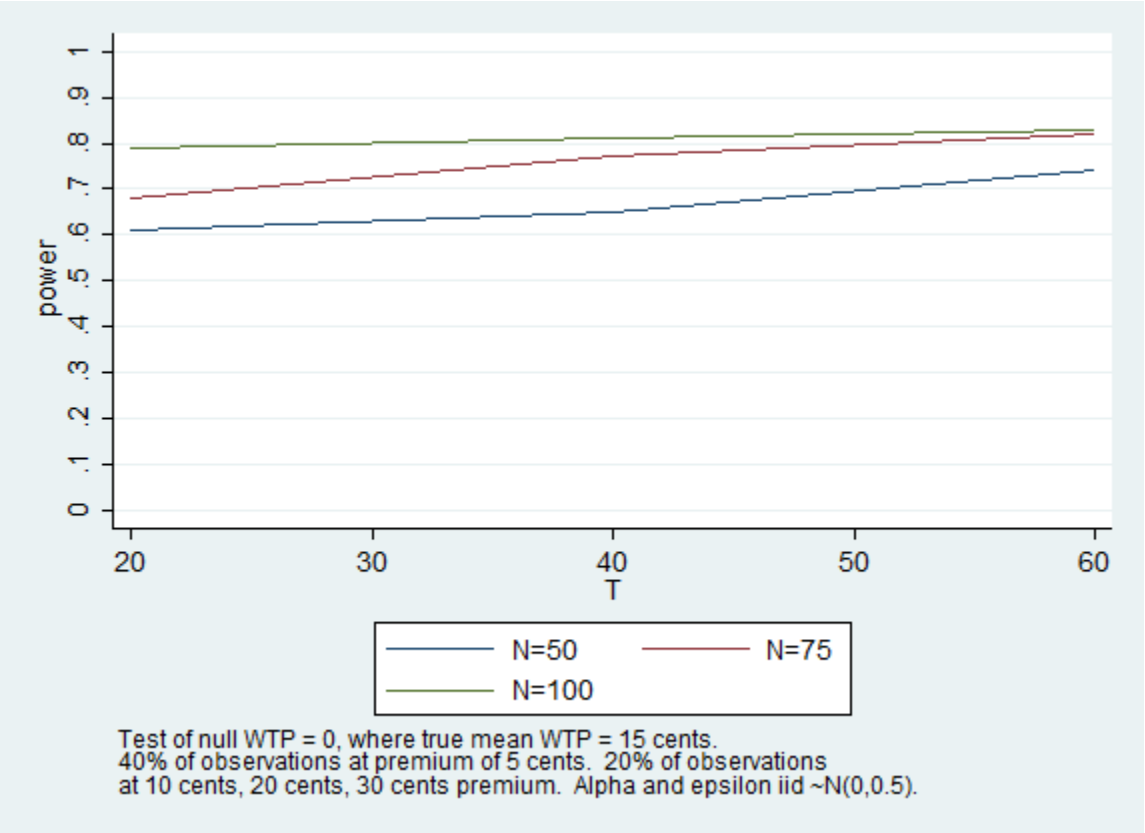
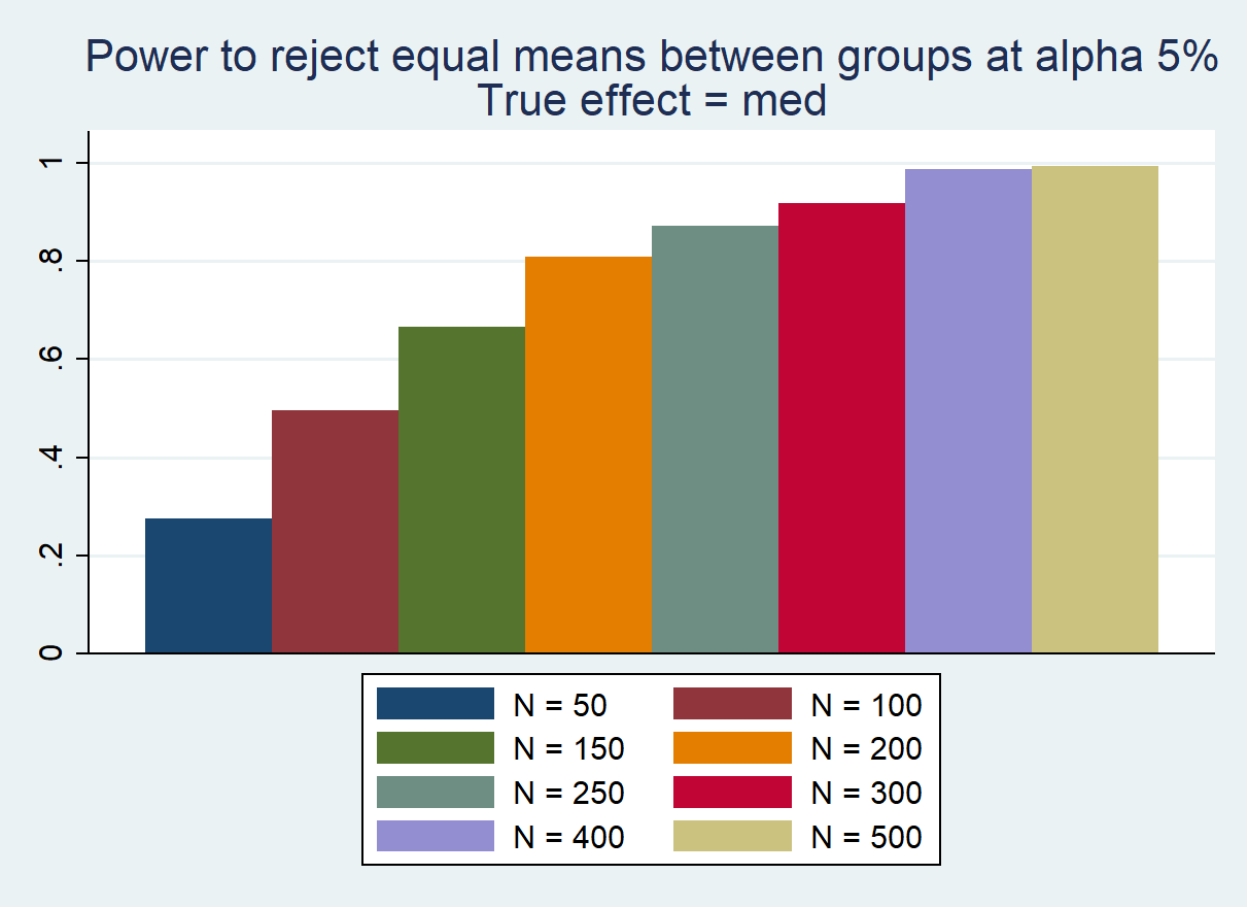


Figure 8 shows a subset of our power calculations by simulation for the IAT in sample B. These were carried out using publicly available data from a standard IAT, the Project Implicit race IAT. We used this data to simulate between-group differences of varying effect sizes, with a range of sample sizes. Power to detect a "medium" effect size of 0.5 SD is over 80% at the planned sample size of 300 respondents completing the IAT.

Figure 8 – Power calculations for Sample B; medium effect size (0.5 SD) on IAT score



9. Data Collection

9.1 Quantitative Instruments

Because the data collection design is central to the evaluation design, we have incorporated discussion of the instruments in Section 8, above. The full instruments are incorporated as an appendix.

9.2 Management of Data Quality

The Premise platform includes several data quality assurance protocols:

- Riders have to take a photograph of their check-in and check-out station;
- The app automatically collects geo-tags and timestamps when a participant enters, records a ride observation, and exits the train;
- Participants have to record the carriage number of the train car on which they ride and submit a picture capturing the number;

- The app also included checks against users changing the time settings on their phone to game the system;
- Only data that is submitted in the right order, at the right time and which represent a credible trip (e.g. check and check-out cannot be done at the same location) are accepted.

Premise provides feedback to riders regarding any issues with the data to improve contributions over time. Data on congestion and male-compliance is collected repeatedly from different sources which allows us to triangulate the data. The platform survey and IAT data collection will be done by only a small team of enumerators directly supervised by the DIME field coordinator.

9.3 Ethical Issues

This IE is subject to review by Duke University's IRB, which has reviewed and approved the study design and informed consent procedures (protocol #D0190). The full protocol and amendments are attached. In Phase 1 (completed), we find that all participants use a mixture of women's and mixed cars, suggesting we are not inducing any woman to ride the mixed car who would not consider doing so. In addition, every individual ride is always completely voluntary so women can choose not to participate on any day, including deciding not to participate in mixed car rides if they prefer (this is an additional margin of adjustment which we measure). In case she experienced harassment, the app directs her to resources available in the Rio area.

10. Data Processing and Analysis

To quantify willingness to pay (Phase 2 – revealed preference experiment), we will estimate:

$$W_{it} = \beta_1 + \beta_2 P_{it} + \beta_3 X_i + \beta_4 Z_t + e_{it}$$

Where W is an indicator for whether woman i chose the women's car on ride t; P is the price (opportunity cost) imposed for the women's car on that ride; X is a vector of individual characteristics, and Z is a vector of ride characteristics. Standard errors are clustered by user. In estimation without control variables, β_2 can be interpreted as the proportion of users who choose to ride the women's car at premium level *P.*

To quantify the impact of riding the women's car on wellbeing (Phase 3 – randomized assignment), we will estimate:

$$Y_{it} = \beta_1 + \beta_2 W_{it} + \beta_3 X_i + \beta_4 Z_t + e_{it}$$

Where Y is an outcome variable (as listed in rows 4 and 5 of Table 1), W is an indicator for whether woman i was assigned to the women's car on ride t, and the remaining variables are as defined above. β_2 is the Average effect of Treatment on the Treated of the riding the women's car for at least part of the ride.

The key groups for heterogeneity analysis will be high/low compliance area-time cells (in which the women's-only rule is being followed), high/low congestion area-time cells (in which crowding on the train is high) and individuals with high / low perceived risk of harassment. We will also report interaction effects with the demographic variables we collect.

The experimental conditions (opportunity cost of women's car in phase 2 and assignment to women's car or mixed car in phase 3) are varied within participant from ride to ride, so balance on participant characteristics is not a concern. We will test for balance on ride characteristics between the experimental conditions.

To address selective participation we will compare the characteristics of participants (in terms of demographics, Supervia use, use of the women's car and perceived risks of harassment) to a representative platform survey, for which takeup in pilot stages is very high (over 90%). The platform survey questionnaire is included in the appendix. Because participation in every stage of the experiment is voluntary, high attrition rates are an issue; we will track attrition between the initial recruitment and subsequent phases of the experiment, and use this to bound our key estimates.

The study will be registered in the RIDIE database after Sample A fieldwork but before Sample B fieldwork. The full set of outcome variables is available in the Appendix and will be included in the RIDIE registry. The Sample A experiment incorporates a limited set of experimental conditions and collection of only a limited set of outcome variables, all of which will be reported in the analysis to address the concern of specification search. We will take steps to address multiple testing in the case of groups of similar outcomes, particularly in the case of multiple measures of psychological wellbeing. We will follow Anderson (2008) and (1) report FDR-adjusted p-values, (2) use a free step-down resampling method to adjust our p-values for familywise error rate, and (3) compute indices of the different measures to test our hypothesis.

11. Study Limitations and Risks

The main limitations of the study are as follows:

1. Participants are self-selected individuals who have a smartphone and respond to the Premise ad. To address this, we avoid mentioning the women's car or any gender issues in all advertising, so that this does not attract individuals who have particularly strong views on this topic. We also use the platform survey to benchmark the sample against a representative sample of Supervia riders.

2. The study examines costs for those who do choose to use the Supervia at rush hour, and not at those for individuals who avoid riding altogether or ride in the middle of the day (as outlined in the Theory of Change). Other literature speaks more directly to these issues. Unlike the existing literature, however, our study speaks more directly to specific mechanisms of gender-based

violence, and addresses the issue of less extreme GBV which may not exclude many women from the public space, but may be more pervasive and impose costs on women when they do go out.

3. The study setting does not allow us to examine spillovers due to displacement of harassment to the mixed car. This does not affect the main analysis of the costs of harassment, because that relies only on a change in the probability of harassment within the women's car to identify this cost. However, it does have important implications for the women's-only car policy itself. Measurement of these spillovers is beyond the scope of the main experimental components of the study, but we will attempt to capture them in focus groups and the platform survey.

REFERENCES

Aguilar, A., Gutierrez, E., & Villagran, P. S. (working paper). Gender Segregation in Public Transportation, Sexual Harassment, and Physical Violence: Evidence from Mexico City's Subway System. Mimeo, ITAM.

Anderson, M. (2008). Multiple inference and gender differences in the effects of early intervention: A reevaluation of the abecedarian, perry preschool, and early training projects. Journal of the American Statistical Association 103 (484), 1481_1495.

Cheema, A., A. Khwaja, F. Naseer, and J. Shapiro (working paper). Crossing Boundaries: Experimental Evidence on Skills Acquisition Access Constraints for Rural Women in Pakistan. Mimeo, Harvard Kennedy School.

Datafolha (2015) http://media.folha.uol.com.br/datafolha/2015/11/09/assedio_sexual.pdf

Duvvury, N., Callan, A., Carney, P. & Raghavendra, S. (2013). Intimate Partner Violence: Economic Costs and Implications for Growth and Development. Women's Voice, Agency & Participation Research Series. Washington, D.C., World Bank

Anthony G. Greenwald, Mahzarin Banaji, and Brian A. Nosek. Understanding and Using the Implicit Association Test: I. An Improved Scoring Algorithm. Journal of Personality and Social Psychology, 85(2):197–216, 2003

Jacoby, H. G., & Mansuri, G. (2011). Crossing boundaries: gender, caste and schooling in rural Pakistan. World Bank Policy Research Working Paper Series.

Jayachandran, S. (2015). The Roots of Gender Inequality in Developing Countries. Annual Review of Economics, (7), 63–88. https://doi.org/10.1146/annurev-economics-080614-115404

Kiplesund, S., & Morton, M. (2014). Gender-Based Violence Prevention: Lessons from World Bank Impact Evaluations. World Bank enGENDER IMPACT, 1–7.

McCollister, K., French, M., & Fang, H. (2010). The Cost of Crime to Society: New Crime-Specific Estimates for Policy and Program Evaluation. Drug Alcohol Depend, 108, 98–109.

Muralidharan, K. and N. Prakash (2013): "Cycling to School: Increasing Secondary School Enrollment for Girls in India," Working Paper 19305, National Bureau of Economic Research.

Velásquez, A. (2014). The economic burden of crime: evidence from Mexico. Mimeo, Duke University.