Benefits and Unintended Consequences of Gender Segregation in Public Transportation: Evidence from Mexico City's Subway System[±]

Arturo Aguilar[†]
ITAM, CIE.

Emilio Gutiérrez
ITAM, CIE.

Paula Soto Villagrán
UAM, Department of Sociology

[±] This project was funded by the Inter-American Development Bank. We are grateful to Carlos Castro for his participation and insights for the study's design. Francisco del Villar and Zeky Murra provided excellent research assistance. Aguilar and Gutiérrez acknowledge support from the Asociación Mexicana de Cultura. All errors are our own.

[†] Corresponding author: Arturo Aguilar, email: arturo.aguilar@itam.mx ; ITAM, Camino a Santa Teresa 930, Col. Heroes de Padierna, 10700 Mexico City, Mexico; Ph. +52-55-56284000

Benefits and Unintended Consequences of Gender Segregation in Public Transportation: Evidence from Mexico City's Subway System

Abstract

For most people, transportation is a basic everyday activity. Costs imposed by violence in the public transportation context might have far reaching economic and social consequences. We conduct a survey and exploit the discontinuity in the hours of operation of a program that reserves subway cars exclusively for women in Mexico City to estimate its impact on self-reported sexual harassment. We find that the program seems to achieve its purpose: it reduces the incidence of sexual harassment towards women by 2.9 percentage points. However, it also produces unintended consequences by increasing the incidence of non-sexual aggression (e.g. insults, shoving) experienced by men in 15.3 percentage points. A willingness-to-pay exercise suggests that from a welfare perspective, it is unclear if the program conveys positive results.

Keywords: Sexual Harassment, Violence, Gender Segregation, Public Transportation, Regression Discontinuity

I. Introduction

The lack of safety in public spaces imposes important costs to individuals, with women being a particularly vulnerable group. The avoidance behavior related to sexual and non-sexual violence may have important consequences in terms of commuting times and choices, labor supply decisions, and access to certain services. According to a study conducted by GfK commissioned by Stop Street Harassment (SSH)[1], 65 percent of women and 25 percent of men in the US report having experienced some kind of sexual harassment[2] in public spaces during their lifetimes. Furthermore, existing evidence shows that women are more responsive to the risk of victimization (Ferraro, 1995), suggesting that unsafe environments impose larger costs on them. While sexual harassment can occur in a variety of public spaces, women's lack of safety in public transportation systems has attracted considerable attention from international organizations and policy makers alike. As a response to this problem, policies that segregate men and women in public transportation spaces have been adopted in several locations, including Mexico, Japan, Brazil, Egypt, Russia, India, Dubai and Iran. Nonetheless, very limited evidence is available to understand the potential effects conveyed by policy interventions that focus on gender segregation.

This paper is the first to present rigorous evidence related to the impact of a gender-segregation policy implemented in a public transportation system on sexual harassment and non-sexual aggression. By using sharp changes in the availability of women-only cars in the Mexico City subway system, we analyze the effects on sexual-related violence (which is the aim of the program), but also study the (possibly unintended) consequences on non-sexual violence, that result from interaction in single-gender environments versus gender-mixed ones. Finally, we perform willingness-to-pay (WTP) exercise to assess the effects on wellbeing on men and women that result from this division. The closest to our study is Kondylis et al. (2016), which measures the impact on

[1] http://www.stopstreetharassment.org/wp-content/uploads/2012/08/2014-National-SSH-Street-Harassment-Report.pdf
[2] Throughout the text, we use the terms "sexual harassment" and "sexual violence" indistinctly.

sexual harassment and willingness to pay to ride women-only subway cars for a sample of 300 women reporting their subway riding experiences through their smartphones in Rio de Janeiro. Our study is conducted from a considerably larger sample (3,466 women), looks at outcomes other than sexual harassment, and also reports previously unexplored impacts of policies of this kind on men's riding experiences, whom are indirectly affected by this policy.[3]

Given the unavailability of existing data sources to conduct the analysis presented in this paper, we collected an on-site survey to subway users around the time at which the program changes its status. The survey was designed specifically to gather evidence about sexual and non-sexual victimization, learn about users' awareness about the program, and their WTP for safety in the subway system. It was sketched with the objective of taking less than 5 minutes to respond.

Our empirical strategy exploits the discontinuity in the program's hours of operation, by comparing self-reported prevalence of sexual and non-sexual violence experienced by subway users before and after it starts operating within each day. After showing that the implementation of the program responds sharply to the mandated schedule, we find that while sexual violence experienced by women decreases, non-sexual aggression suffered by men increases. Through a WTP exercise we find that sexual violence imposes considerably large costs to subway users, but non-sexual aggression also generates non-negligible costs. Taken together, our results then suggest that, while actually having benefits on reducing sexual violence (as the program intended), the gender segregation program implemented in Mexico City also produces costly consequences on male users, making it unclear from a welfare perspective if the overall result of the program is positive.

[3] Kondylis et al (2016) compare the prevalence on sexual harassment in women-only versus mixed subway cars in a context in which the program is always operating. However, the gender composition and congestion levels of mixed cars (as shown in this paper) may also be affected by the women-only cars policy. Precisely trying to estimate indirect impacts of the program on mixed cars, our paper compares the outcomes of interest during times of the day when the program is and is not in operation.

The existing literature exploring the relationship between gender and transportation has generally focused on quantifying gender gaps in travel times and transport-mode choices, and sketching theories rationalizing how the gender differences in household chores, occupation and/or labor force participation may explain them (see Boarnet and Hsu (2015) for a review of this literature).[4] Within this literature, some work identifies important gender differences in victimization and fear of violence or crime, and suggests them as the potential forces between some of the gender differences in transport choice (Lynch and Atkins; 1988; Ferraro, 1995).[5]

While gender differences in victimization and fear of crime in public spaces has been documented, only recently have international organizations, media outlets and policy makers increased their attention on discussing and designing policies and interventions aimed at reducing gender-based violence in public spaces.[6] To our knowledge, ours is the first paper to analyze how gender segregation in the public transport system can heterogeneously influence violent interactions and to quantitatively evaluate the effects of this type of policies that are recently gaining popularity. Apart from Kondylis et al (2016), the closest existing work is a paper that finds that women are willing to sacrifice educational quality in order to avoid risky travel routes (Borker 2017) and ethnographic analyses that look into gender interactions and responses (Dunckel-Graglia, 2013; Agrawal and Sharma, 2015).

[4] For specific papers documenting differences in commuting times by gender see Ericksen (1977) Hanson and Johnston (1985) and Johnston-Anumonwo (1997). For the evolution of these differences with respect to female labor force participation see Crane (2007) and Crane and Takahashi (2009). With respect to differences in travel times for non-work related activities see Hanson and Hanson (1981), Mauch and Taylor (1997), Handy (1998),Steiner (2000), Sarmiento (1998), McGuckin and Murakami (1999), and Hjorthol (2000).

[5] Because we explore the effects that separating men from women in public spaces, our paper may also be informative to the existing literature on segregation and conflict (see, for example, Cutler and Glaeser, 1995 and Flaherty and Sethi, 2010), and on the effects of single gender vs mixed-gender schools (see, for instance, Martin and Fabes, 2001, and Halper et al, 2011).

[6] For instance, in 2011, the United Nations Committee to for the Status of Women launched the UN Women's Safe Cities Global Initiative, with the participation of 14 cities (including Mexico City) from an equal number of countries around the world.

The remainder of the paper is organized as follows. Section II describes the context and the program implemented in Mexico City's subway system. Section III describes the survey instrument and Section IV the empirical strategy. Section V presents the results that measure the impact of the program on sexual and non-sexual violence, with a series of robustness checks in Section VI. In section VII we present out WTP analysis, and the last section concludes.

II. Context description and Mexico City's "*Women Travelling Safely*" program

Mexico City and its subway system are not exempt from gender-based violence.[7] The city hosts a large public transportation system, generally characterized by heavy congestion, which responds to the needs of over 20 million inhabitants.[8] Between October and December 2014, the Mexico City subway system alone took charge of 4.46 million trips on average per day (18 percent of all the trips in the public transportation system). While ranked highest among the alternatives for public transportation in the city in terms of quality,[9] the existing statistics regarding the prevalence of violence, and particularly sexual violence, are staggering.

Mexico City's public transportation system is ranked the second least safe among the fifteen largest cities in the world,[10] and the least safe in Latin America.[11] According to a survey (Garibi et al. 2010) conducted by the National Council for the Prevention of Discrimination (*Consejo Nacional para Prevenir la Discriminación, CONAPRED*) in three of the main public transportation hubs in the city, more than 80 percent of female passengers were victims of some kind of sexual violence in the public transportation system during 2009 (63.9 percent received unwanted sexual looks and 37.7

[7] On April 25, 2016, a massive demonstration was held in Mexico City, organized by women demanding that the government design and implement policies to fight sexual violence against them in public places. http://www.eluniversal.com.mx/articulo/metropoli/df/2015/11/25/marcha-contra-la-violencia-hacia-la-mujer-llega-al-zocalo

[8] http://www2.inecc.gob.mx/publicaciones/libros/652/vallemexico.pdf

[9] http://www.parametria.com.mx/carta_parametrica.php?cp=4539

[10] http://mx.reuters.com/article/topNews/idMXL2N0SJ0N320141029?sp=true

[11] IADB (2015) https://publications.iadb.org/bitstream/handle/11319/7441/El-porque-de-la-relacion-entre-genero-y-transporte.PDF?sequence=4

percent declared having been inappropriately touched in the previous year), and more than 80 percent of those aggressions took place in the subway system.

Given the attention that this type of violence has received by international organizations and policy makers around the world, and its prevalence in Mexico City, in 2008 government officials launched a multi-level intervention named "*Viajemos Seguras*" (*Women Travelling Safely*). The program comprises legal support for women who suffered any kind of violence during their commute, and the availability of women-only public transportation options, such as women-only passenger cars[12] in various subway lines, women-only sections in the rapid bus transit system, women-only buses, and even women-only taxis. In the subway, the program reserves the front three cars of the train for women during the times of the program's operation: before 10:00 A.M. and after 2:00 P.M. Enforcement of the program is managed by police officers present in subway stations that use different means to limit access to men, ranging between a tape in the floor that signals the three car limit to gates regulated by officers. Men that ignore the decree could be detained and fined.[13] Qualitative work performed complementarily to the quantitative analysis presented in this paper suggests that transgressors are rarely prosecuted. Nonetheless, it was repeatedly observed that men that travel in a women-only car tend to be heavily punished socially (e.g. women confront and yell at them, report them to officers, and demand them to abandon the car).

While clear in its objectives, to our knowledge, none of the "*Viajemos Seguras*" program's components has been rigorously evaluated, and very little discussion regarding its potential effects on outcomes other than sexual violence has taken place.[14] To our knowledge, this paper is the first to

[12] Children are also allowed to ride women-only passenger cars.
[13] Regulated under the Law of Civic Culture of the Federal District. Article 26, numeral XI indicates that transgressors could serve between 25 to 36 hours of detention and fined between 1,505 and 2,150 Mexican Pesos (approximately between 80 and 115 USD). However, transgressor are seldom prosecuted.
[14] Some groups have criticized programs like this given that they lack a strategy to combat the problem's source. Some have discussed the potential costs of these programs in terms of agglomeration. See for example, http://www.news.com.au/travel/travel-updates/womenonly-pink-carriages-idea-for-aussie-trains-causes-controversy/news-story/8377482b8b705dd2854a51d0eb0b7847

empirically measure the impact of the designation of women-only subway cars on both sexual and non-sexual violence, as experienced by both men and women in Mexico City's public transportation system.

III. Survey Design and Implementation, and Descriptive Statistics

The evaluation of a program of this kind is an empirical challenge due to obvious concerns of reverse causality, and the unavailability of information regarding the prevalence of any kind of violence in the subway system. To overcome these limitations we took advantage of the precise schedule for the program's implementation, and administered a survey designed ad-hoc to the information required to 5,000 individuals (1,534 men and 3,466 women) with a higher emphasis around the times when the program stops and restarts. The survey employed was short in design and gathered from subway users with the purpose of obtaining information about the program's implementation, the incidence of different kinds of sexual and non-sexual violence, and a short willingness-to-pay for safe travel questionnaire. Because of the sensitive nature of the questions asked, the eight surveyors recruited were all female, graduate students in social sciences-related fields at one of the largest public universities in the city (Universidad Autónoma Metropolitana, UAM), with experience in having conducted surveys in the past and familiar to the gender-violence issues surrounding Mexico City's public transportation system.[15]

The key aspect to identify the program's impact is the short term and sudden variation that occurs at the time in which the program is suspended (10:00 A.M.) and then restarted (2:00 P.M.). In order to be able to capture short term variations in both the implementation of the program and the prevalence of sexual and non-sexual violence, surveyors stood at subway stations waiting for arriving trains (mainly around the time in which the program changed its status), and were trained to recruit respondents ensuring that the survey duration was short enough for them to be able to finish it while

[15] Surveyors wore a tag and identified themselves as members of UAM and asked for the respondents consent before the implementation of the survey.

walking out of the subway station. These constraints required the design of a very short questionnaire and imposed important limitations to the information that could be obtained. However, after a series of pilot surveys (during which the surveyors also gained experience conducting it), we reached a final questionnaire that took, on average, five minutes to complete. In the appendix, we provide the full content of the questionnaire used.

The survey consisted of three main sections with 25 questions in total. The first section recovers socio-economic characteristics of respondents (age, marital status, number of children, labor force participation, car ownership and subway ridership frequency). The second section refers specifically to the subway trip that the subjects had just taken. It recovers the subway station where their trip started, a question explicitly asking whether the women-only cars program was enforced during the trip, a series of questions regarding whether the respondent was a victim or a witness of any kind of sexual (unwanted sexual looks, sexually-charged comments, pictures of them taken without consent, inappropriate touching, brushing against their body, and other types of sexual aggression) or non-sexual violence (pushing/shoving, physical fights, insults, theft) in addition to the aggressor and the victim's gender, among others.

Finally, the last section recovers the response to two dichotomous choice questions aimed at recovering the subjects' WTP for safety in the subway system in general, and during the last trip taken in particular. The first question stated: "Given your experience on TODAY's trip, would you prefer to pay X Mexican pesos for a completely safe trip?"[16], and the second question stated: "Would you be willing to ALWAYS pay X pesos if in exchange the subway system guaranteed your safety during your trip?" In both cases, the price X was randomly allocated in different surveys and could either be equal to six, eight or ten Mexican pesos. A total of 5,000 surveys were successfully

[16] The price of a subway ride in Mexico City is six pesos. The alternative amounts stated in the WTP exercise were 6, 8 and 10 pesos, randomly assigned across questionnaires.

administered on work days from February to December 14, 2014, of which 30.7% of respondents were male and the remaining 69.3% were female.

Descriptive statistics of socio-economic variables for women and men are presented in Table 1, columns 1 and 2, respectively. In order to better assess the representativeness of the characteristics of subway users surveyed, columns 3 and 4 show mean levels for the same set of descriptive statistics according to CONAPRED's survey conducted on public transportation users in 2009 (Garibi et al., 2010). As can be seen, the age distribution of our respondents does not differ significantly from that of the individuals surveyed by CONAPRED. Close to fifty percent of respondents are aged between 15 and 29 years old, and close to one quarter between 30 and 44. Respondents in our survey have higher schooling levels: 42 percent of women and 38 percent of men report having completed college, while only 11 percent of women and 17 percent of men in the CONAPRED survey have a college degree. Our respondents also use the public transportation system less frequently: 22 percent of women and 18 percent of men in our survey use the public transportation system daily, while 54 and 61 percent of those surveyed by CONAPRED report to be daily users.

While these differences may imply external validity concerns, to the extent that our empirical strategy correctly exploits quasi-random variation in the program's implementation, our results can be interpreted as causal for individuals riding the subway around the times at which the program starts and stop operating.

Table 1

Descriptive Statistics				
	Survey		CONAPRED**	
	Women	Men	Women	Men
Married*	25.7	31.2		
Has children*	46.4	46.4		
Num of children	1.0	1.1		
Age*				
15-29	54.7	49.7	53.6	52.2
30-44	26.0	28.8	30.0	34.1
45-59	14.3	15.0	14.6	11.7
60+	4.4	5.9	1.9	2.0
Schooling*				
Less than Primary	1.4	1.0	2.6	0.3
Primary	5.4	4.5	14.6	11.0
Secondary	13.2	15.1	37.8	33.1
High School	30.6	34.8	32.6	37.5
College	42.2	38.3	11.2	17.4
Graduate	5.1	5.3	0.0	0.3
Number of Days a Weeks using Public Transportation*				
Everyday	22.3	18.2	53.6	61.2
Mon-Fri	41.6	40.1	35.6	33.1
Other	36.1	41.7	10.8	5.7
Observations	3466	1534	267	299

* Reported values in percentages

** Source: Garibi et al. (2010)

Table 2 shows descriptive statistics regarding the prevalence of sexual and non-sexual violence experienced and witnessed by our survey respondents. A few facts are worth highlighting. First, consistent with evidence for other contexts, women are considerably more likely to experience sexual violence: 7.3 percent of all surveyed women and 2.3 percent of surveyed men report having been victims of some kind of sexual violence during their last trip. Second, while unwanted sexual looks are the most prevalent form of sexual violence (5.9 percent of women report having received them), other forms of sexual violence are also prevalent (1.2 and 1.3 percent of surveyed women report brushing against their bodies and receiving sexually-charged comments, respectively). Third, within non-sexual violence, pushing and shoving is the most frequent category being reported, with 22.3 percent of surveyed women and 18.3 percent of men experiencing this type of violence. Fourth, observed violence is in general higher than experienced violence. Fifth, the ranking with respect to incidence for different categories of sexual and non-sexual violence is the same for men and women, and women report being victims of either type of violence more frequently.

Table 2

Descriptive Statistics Prevalence of Sexual and Physical Violence*				
	Experienced		Observed	
	Women	Men	Women	Men
Non-sexual violence	**22.9**	**18.3**	**32.9**	**27.9**
Pushing / shoving	22.3	17.9	31.3	26.1
Fights	0.8	0.7	2.4	2.5
Insults	1.6	1.3	7.1	5.8
Theft	0.3	0.3	0.7	0.9
Sexual violence	**7.3**	**2.3**	**9.8**	**10.8**
Unwanted looks	5.9	1.1	8.3	8.3
Sexual comments	1.3	0.4	1.3	1.8
Pictures taken	0.2	0.2	0.5	0.7
Inappropiate touching	0.4	0.4	1.3	1.8
Brushing against body	1.2	0.5	2.5	2.5
Observations	3,466	1,534	3,466	1,534

* Reported values in percentages

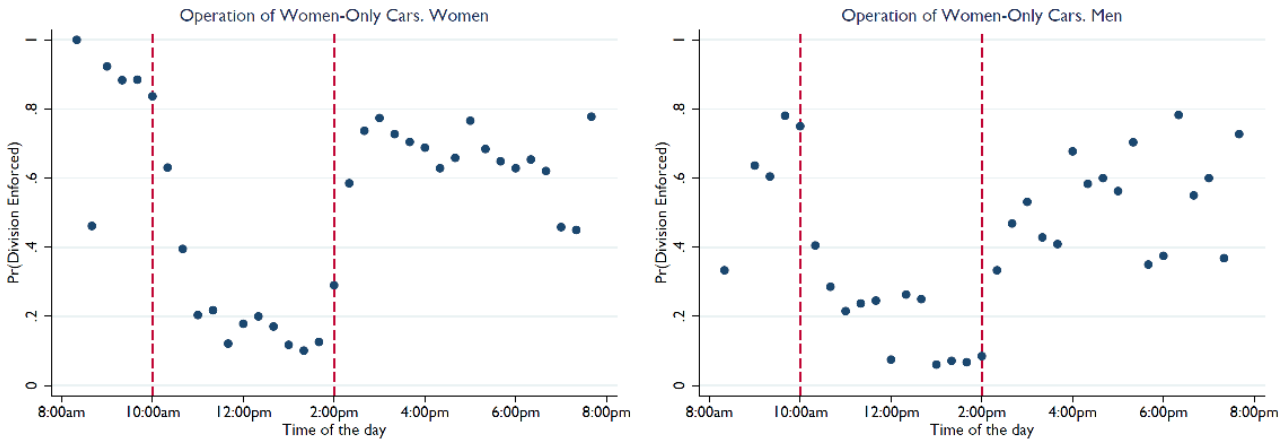
IV. Empirical Strategy

The program's implementation rules mandate its implementation in two periods each day: first from 6:00 A.M. to 10:00 A.M. and second from 2:00 P.M. to 8:00 P.M. Our empirical strategy exploits the sharp change in the program's operation at 10:00 A.M. and 2:00 P.M. and then compares self-reported measures of experienced sexual and non-sexual violence during the surveyed individuals' last trip through a regression discontinuity design.

Showing that the program's implementation responds sharply to the mandated hours of operation is crucial for identification. Figure 1 explores graphically if this was indeed the case. It plots the fraction of women and men that self-reported that the program was in operation versus the time at which they boarded the subway (in 5-minute intervals). As can be seen, there is a clear sharp

decrease in the fraction of people declaring that the women-only program operated during their last trip at 10:00 A.M., and a sharp increase in the same variable at 2:00 P.M. Both jumps in the program operation variable correspond to the mandated hours of operation of the program.

Figure 1
Program's Operation



Note: Estimations done by the authors. Each point represents a 20-minute interval mean.

In order to explore differences in the program's implementation (and impact) between the morning and afternoon, and to explore differences by gender, we define four different samples for the empirical analysis: two for women and two for men, each consisting of individuals of each gender that boarded the subway before and after noon. For those boarding the subway before noon, the discontinuity of interest will be 10:00 A.M., while for those boarding after noon, we will exploit the discontinuity observed in the implementation of the program at 2:00 P.M. For each sample, the specification will be the following:

$$D_{idt} = \gamma_d + f(H_{it} - h) + \beta T_{it} + \alpha X_i + \varepsilon_{idt} \qquad (1)$$

Where D_{idt} is a dummy variable taking value of one if individual *i*, on day *d*, observed the women-only subway cars program being implemented when (s)he boarded the subway at time *t*; γ_d are day

fixed-effects; $f(.)$ is a smooth non-parametric function; H_{it}, represents the time at which the survey was conducted; h, is employed to center the smooth function at the time in which the program changes its status (for the morning sample it is 10:00 A.M. and for the afternoon it is 2:00 P.M.); T_{it}, is a dummy taking value of one if the survey was conducted at a time of the day at which, according to the regulation, the program should be implemented (before 10:00 A.M. for the morning samples and after 2:00 P.M. for the afternoon); X_i, is a vector of individual characteristics; and ε_{idt} is an error term. In particular, as most of the recent papers that exploit regression discontinuity designs, we implement the Calonico et al (2014)'s non-parametric robust regression discontinuity command in Stata in our main specifications. We also present results pooling the morning and afternoon samples for each gender.

In order to explore if the program had an effect both on sexual and non-sexual violence, we run the same specifications, using a dummy variable taking value of one if the individual experienced any kind of sexual and non-sexual violence during the last subway trip as dependent variables. The results of these regressions can be understood as a reduced-form estimate of the impact of the program on different kinds of violence. By design, it is also possible to use an instrumental variable estimate of the impact of the program using the program's implementation time as instrument. Because we worry that the program's implementation may be measured with error, and that this error may be negatively correlated with the forcing variable (distance with respect to the cutoff), throughout the main text we report the reduced-form estimates.

V. Results

a. Program operation

Table 3 shows the estimation results for equation (1), using the dummy variable indicating whether the respondent declared that the women-only cars program was operating when (s)he boarded the subway as dependent variable, and adjusting for a non-parametric flexible relationship between the

time at which the survey was conducted and the outcome of interest. Columns 1 and 4 correspond to the estimation of equation (1), pooling the morning and afternoon samples for women and men, respectively. Columns 2 and 5 report the results of estimating equation (1) during the morning sample (i.e. before 12:00 P.M.) for the women's and men's samples, respectively. Columns 3 and 6 report results of estimating equation (1) for the women's and men's afternoon samples (i.e. after 12:00 P.M.), respectively.

Consistent with the evidence presented in Figure 1, our results indicate that the mandated hours of operation of the program strongly predict its implementation. For the pooled samples, individuals that boarded the subway are close to 20 percentage points more likely to report the program being implemented at the discontinuities (17 and 23 percentage points for women and men, respectively).

Table 3

First stage regression results
Dependent Variable: Dummy=1 if program was enforced

	Women			Men		
	Pooled	Morning	Afternoon	Pooled	Morning	Afternoon
Program Scheduled to Operate	0.172***	0.055	0.421***	0.232***	0.288***	0.202***
	[0.0292]	[0.0339]	[0.0511]	[0.0518]	[0.0678]	[0.0642]
Observations	3,466	1,831	1,635	1,534	776	758
Controls						
Date FE	Yes	Yes	Yes	Yes	Yes	Yes
Socio-Economic variables	Yes	Yes	Yes	Yes	Yes	Yes

Specification using a local polynomial with triangular kernel and 120 minutes bandwidth
Robust standard errors in brackets
* significant at 10%, ** significant at 5%, *** significant at 1%

b. Effects on sexual violence

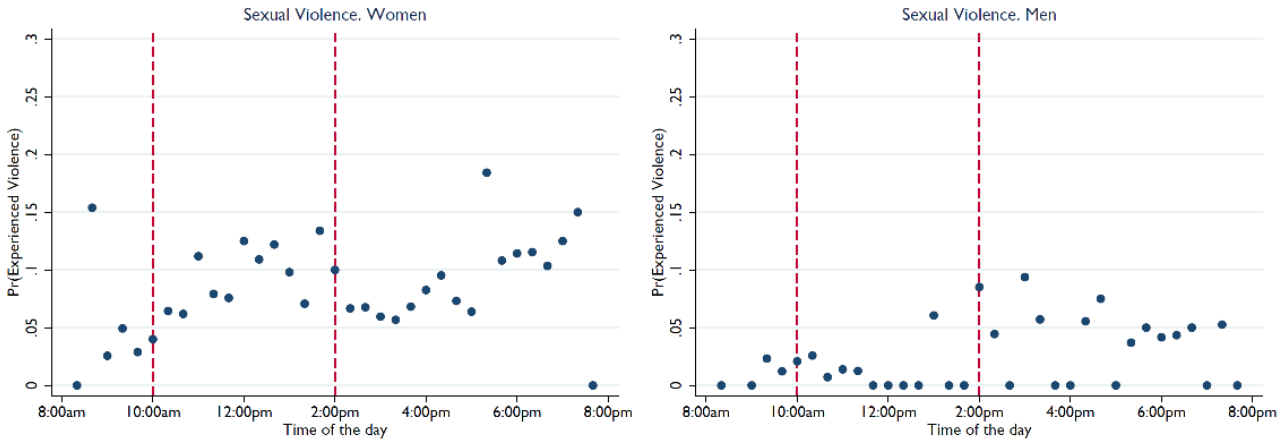
To explore the potential effects of the program on sexual and non-sexual violence we begin by exploring the conditional means of the prevalence of both types of violence with respect to the time of the day in which the individual surveyed boarded the subway. We pay special attention to the times in which the program's operation sharply changes.

Figure 2 shows the prevalence of sexual violence, whose reduction is the main purpose of the program. It displays the levels of sexual violence by gender of the individual interviewed. As described in Table 2, sexual violence could result from the occurrence of unwanted looks, sexually charged comments, pictures taken without consent, inappropriate touching or brushing against the victim's body.

Figure 2 shows that men rarely suffer from sexual violence; the few positive reports concentrate in the afternoon. In terms of women, no effect seems to happen around the morning cutoff and the afternoon displays some evidence of a reduction in sexual violence.

Figure 2

Incidence of sexual violence



Note: Estimations done by the authors. Each point represents a 20-minute interval mean.

Table 4 confirms the results observed in the graphs through our regression analysis. All columns correspond to the same sample as the corresponding columns in Table 3. The program displays some evidence of reducing the incidence of sexual violence and, as expected, the effect is only significant for women.[17] For the pooled (afternoon) sample, women are 2.9 (5.3) percentage points less likely to report having been victims of any kind of sexual violence at the discontinuity. Given the sample averages presented in Table 2, the effects are considerably large: the incidence of sexual violence is

[17] Complementary information obtained from qualitative work suggests that women diminish their perception of insecurity thanks to the accessibility to women-only cars and that they agree with the implementation of the program.

reduced by 39 percent. Despite not being significant at the traditional levels, the afternoon effect seems to drive the effects and displays a p-value marginally below 0.1 (p-value=0.12). The next section shows robustness checks parametrizing the relationship between the distance (in time) to the program's operation through different order polynomials. The results consistently show a negative effect mainly driven by the afternoon effect.

Table A.1 disaggregates the results of women found in Table 4 by type of sexual aggression. The findings suggest that the effect seems to be mainly driven by "unwanted looks" and "inappropriate touching". Both indicators display significant reductions in the order of 3.1 and 0.8 percentage points, respectively.

Table 4
Effects of the program on sexual violence

Reduced form regression results
Dependent Variable: Dummy=1 if sexual violence was experienced

	Women			Men		
	Pooled	Morning	Afternoon	Pooled	Morning	Afternoon
Program Scheduled to Operate	-0.029*	-0.019	-0.053	0.004	0.0004	0.019
	[0.0156]	[0.0180]	[0.0344]	[0.0197]	[0.0205]	[0.0401]
Observations	3,466	1,831	1,635	1,534	776	758
Controls						
Date FE	Yes	Yes	Yes	Yes	Yes	Yes
Socio-Economic variables	Yes	Yes	Yes	Yes	Yes	Yes

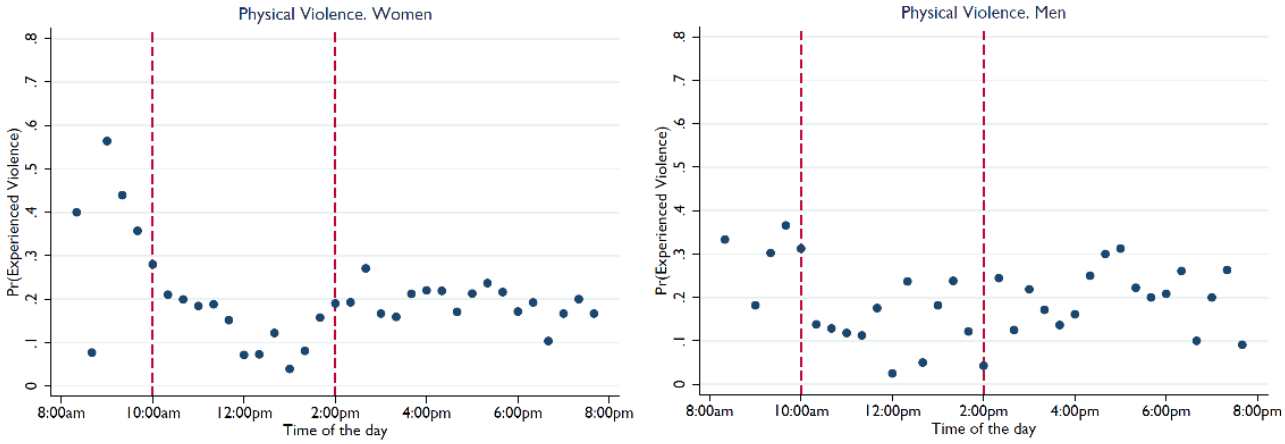
Specification using a local polynomial with triangular kernel and 120 minutes bandwidth
Robust standard errors in brackets
* significant at 10%, ** significant at 5%, *** significant at 1%

c. Effects on non-sexual violence

To analyze the effects on non-sexual violence, Figure 3 begins by presenting the conditional means' graph of this outcome's incidence by the hour of the day in a similar fashion to Figure 2. The graphs suggest that the morning implementation of the program seems to drive up non-sexual violence for both men and women. As for the afternoon, a smaller increase seems to occur for the women sample only.

Figure 3

Incidence of non-sexual violence



Note: Estimations done by the authors. Each point represents a 20-minute interval mean.

Table 5 reports the point estimates obtained from our regression analysis. Perhaps surprisingly, while the program seems to be reducing the incidence of sexual violence against women, it also seems to affect *positively* the incidence of non-sexual violence for men. For the pooled samples, men are 15 percentage points more likely to have experienced any kind of non-sexual violence as a result of the program's implementation. This effect is mainly driven by the morning discontinuity. It is worth noting that while the effect for the female samples is smaller in magnitude and generally not significantly different from zero, the point estimate for the pooled sample is positive and obtains a p-value slightly below 0.1 (p-value=0.16).

Table A.2 in the appendix suggests that the increase in non-sexual violence experienced by men is mainly explained by a 14.6 percentage point increase in "pushing / shoving", although "insults" also displays an increase of 2.1 percentage points (p-value=0.13). As for women, despite not obtaining significant aggregated effects for the non-sexual violence indicator in Table 5, we also detect a lower magnitude increase in "pushing/shoving" of 3.6 percentage points (p-value=0.21) and "insults" of 1.9 percentage points (p-value=0.008), the latter being significant at the traditional levels.

Table 5
Effects of the program on non-sexual violence

Reduced form regression results
Dependent Variable: Dummy=1 if non-sexual violence was experienced

	Women			Men		
	Pooled	Morning	Afternoon	Pooled	Morning	Afternoon
Program Scheduled to Operate	0.040	0.008	-0.011	0.153***	0.232***	0.036
	[0.0289]	[0.0389]	[0.0470]	[0.0435]	[0.0664]	[0.0514]
Observations	3,466	1,831	1,635	1,534	776	758
Controls						
Date FE	Yes	Yes	Yes	Yes	Yes	Yes
Socio-Economic variables	Yes	Yes	Yes	Yes	Yes	Yes

Specification using a local polynomial with triangular kernel and 120 minutes bandwidth
Robust standard errors in brackets
* significant at 10%, ** significant at 5%, *** significant at 1%

Identifying the mechanisms through which the program affected non-sexual violence is unfortunately beyond the scope of this paper. The effect may be driven by differences in agglomeration across subway cars driven by the program's implementation or by differences in social norms regarding aggressive behavior in gender segregated versus unsegregated environments, for example. However, in Appendix 2, we sketch a simple model that illustrates how such an effect can arise even when abstracting from these forces, simply by reducing differences in observable characteristics across subway car passengers.

VI. Robustness checks

Relatively strong assumptions have to hold for the results presented in the previous section to be interpreted as the impact of the program on the incidence of sexual and non-sexual violence. First, as in any RD analysis, we require that only the outcome of interest varies sharply at the discontinuity for our estimates to be interpreted as causal. Second, the functional form chosen to control for the relationship between time of the day and the outcomes of interest should be smooth enough to capture the relationship between these variables in the absence of the program. And third, individuals surveyed must not perfectly manipulate or respond to the known schedule of the program's operation

by changing the time at which they board the subway around the discontinuity. In this section, we provide evidence supporting the validity of these assumptions for the context analyzed.

Table 6 reports means of socio-economic characteristics for women and men divided by two groups: those who boarded the subway at a time when the program is in operation (i.e. before 10:00 A.M. and after 2:00 P.M.) and those who boarded when the program is not available (i.e. between 10:00 A.M. and 2:00 P.M.). Columns 1 to 4 focus on the women's sample, while columns 5 to 8 restrict the sample to men. In particular, column 3 shows mean differences between characteristics of women that boarded the subway when the program was in operation and women that boarded the subway when the program was not active. Column 7 does the same calculation for men. Lastly, columns 4 and 8 show the results for women and men, respectively, of implementing the baseline specification described in equation (1) without including any of the SES variables as controls and using each of the characteristics as the outcome of interest.

Perhaps not surprisingly, women and men who board the subway when the program is in operation differ significantly in observable characteristics from those who board it when it is not. In particular, women who board the subway when the program is not scheduled to operate are less likely to be employed and use less frequently the subway than those who board it when the program is operating. Men who board the subway when the program is not operating are more likely to have children, to have attended college, and less likely to own a car and be frequent subway users. Nonetheless, according to the estimates presented in Columns 4 and 8, there are no significant differences in observable characteristics for both the women and men samples around the discontinuities in the program's operating schedule. This evidence favors the assumption that the change at the threshold can be exploited as exogenous variation in the program's operation.

Table 6

Exogeneity test

Robustness: Test for Smoothness of Other Variables at the Discontinuity								
	Women				Men			
	Between 10:00 A.M. and 2:00 P.M.	Before 10:00 A.M. and after 2:00 P.M.	Difference	RD estimate	Between 10:00 A.M. and 2:00 P.M.	Before 10:00 A.M. and after 2:00 P.M.	Difference	RD estimate
Married	0.26	0.26	-0.001	0.021	0.32	0.31	-0.01	0.088*
			[0.0150]	[0.0304]			[0.0238]	[0.0524]
Has children	0.48	0.46	-0.018	-0.012	0.49	0.44	-0.053**	0.06
			[0.0171]	[0.0363]			[0.0256]	[0.0576]
Age	31.74	31.77	0.034	0.114	33.39	32.82	-0.575	2.71
			[0.450]	[1.000]			[0.695]	[1.616]
High school degree or less	0.51	0.50	-0.008	0.023	0.59	0.51	-0.074***	-0.0001
			[0.0172]	[0.0364]			[0.0255]	[0.0581]
Student	0.42	0.42	0.001	-0.021	0.35	0.39	0.040	-0.086
			[0.0174]	[0.0371]			[0.0252]	[0.0582]
Employed	0.61	0.67	0.060***	-0.009	0.75	0.74	-0.007	0.033
			[0.0167]	[0.0363]			[0.0227]	[0.0530]
Car ownership	0.43	0.41	-0.024	0.02	0.47	0.54	0.073***	0.02
			[0.0172]	[0.0366]			[0.0259]	[0.0582]
Frequent subway ridership (at least five times a week)	0.60	0.67	0.068***	-0.005	0.56	0.61	0.048*	0.007
			[0.0165]	[0.0344]			[0.0253]	[0.0564]
Observations	1,485	1,981		3,466	846	688		1,534

RD estimate specification using a third degree polynomial with different trend in the morning and afternoon
Robust standard errors in brackets
* significant at 10%, ** significant at 5%, *** significant at 1%

Table 7 replicates the main results using as outcomes the dummy variables indicating whether the program was operating, and whether the respondent experienced sexual or non-sexual violence. Each column indicates the result of different specifications and each line presents the results for a different outcome variable. In particular, columns 1 and 5 replicate the non-parametric analysis presented throughout the main text. Columns 2 and 6 control for a linear relationship between the distance (in time) from the discontinuity and the outcome of interest (on each side of the discontinuity). Columns 3 and 7 control for quadratic time trends. And Columns 4 and 8 include a cubic time trends on each side of the discontinuities. Columns 1 through 4 focus on the women sample, while the remaining focus their attention on men.

Table 7
Functional form's test

Robustness: Test for Different Functional Forms								
	Women				Men			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
First Stage:								
Program Operation	0.172***	0.305***	0.273***	0.222***	0.232***	0.260***	0.287***	0.253***
	[0.0292]	[0.0240]	[0.0277]	[0.0305]	[0.0518]	[0.0392]	[0.0472]	[0.0478]
Reduced Form:								
Sexual violence	-0.029*	-0.0484***	-0.0437**	-0.0240	0.004	0.0197	0.0114	0.0110
	[0.0156]	[0.0144]	[0.0180]	[0.0192]	[0.0197]	[0.0147]	[0.0181]	[0.0200]
Non-sexual violence	0.040	0.0597***	0.0359	0.0148	0.153***	0.122***	0.138***	0.122***
	[0.0289]	[0.0225]	[0.0282]	[0.0300]	[0.0435]	[0.0339]	[0.0428]	[0.0438]
Functional form	Kernel regression	Linear	Quadratic	Cubic	Kernel regression	Linear	Quadratic	Cubic
Sample	Pooled	Pooled	Pooled	Pooled	Pooled	Pooled	Pooled	Pooled
Date FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Socio-Economic Variables	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Columns (1) and (5) correspond to our baseline results. Polynomial specifications with different trend in the morning and afternoon
Robust standard errors in brackets
* significant at 10%, ** significant at 5%, *** significant at 1%

According to the results presented using the program's operation as dependent variable, the estimated impact of the discontinuities on the probability that respondents declare having observed the program remains positive, large in magnitude, and statistically different from zero at a high confidence level throughout specifications.

The second line in Table 7 shows results of the parametric regressions using sexual violence as dependent variable. The sign of the coefficients for the change in the prevalence of sexual violence remains negative and significant at the traditional levels in three out of the four specifications for the women's sample, and positive, close to zero and not significant for the men's sample. In fact, the estimate of the negative impact of the program on the prevalence of sexual violence for women is larger than in our main analysis when only controlling for linear and quadratic time trends.

Finally, the third line in Table 7 shows that regardless of the specification, except for the specification that only controls for linear time trends, we do not find statistically significant estimates for the impact of the program on non-sexual violence for the women's sample. For the men's sample, however, the estimated coefficients remain relatively stable with the inclusion of more flexible time

trends, positive and statistically significant, giving us confidence that the women-only subway program had an undesired effect by increasing the prevalence of non-sexual violence among men.

VII. Willingness to Pay

The results presented thus far give suggestive evidence that the women-only subway cars program in Mexico City did achieve its desired effect of reducing sexual violence against women, but unexpectedly increased physical violence against men. While the staggering statistics regarding the prevalence of sexual violence in Mexico City's public transportation system call for policy interventions aimed at reducing it, it may be important to quantify the potential costs that such policies may entail.

A full cost-benefit analysis of the program is beyond the scope of this paper. However, given that our questionnaire did include a short willingness-to-pay exercise, in this section we explore how prices and the incidence of sexual and non-sexual violence correlate with individuals' likelihood of responding affirmatively to the question of whether they would have paid a higher subway price in exchange for having experienced a "completely safe trip". In section 3 we detailed that in the design of this exercise, the "higher price" was randomly allocated, that is, three different versions of the questionnaire were printed and randomly distributed to surveyors. The three versions included different levels of the "higher price": 6, 8 and 10 pesos, which respectively correspond to a 20, 60 and 100 percent increase in the current price, which is 5 pesos.

We begin by estimating a standard OLS regression using a dummy variable taking value of one if individual's response to the WTP question was affirmative as the dependent variable, against the log of the price offered. With this estimation, we analyze if, as expected, respondents are sensible to the price as a standard demand system would predict. Then, following the WTP literature we us a logit estimation and add the self-reported prevalence of physical and non-sexual violence as explanatory variables (in addition to the log of the price).

Specifically we estimate with a logit:

$$\log\left(\frac{W_{idt}=1}{W_{idt}=0}\right) = \gamma_d + \beta_1 \log(price)_i + \beta_2 NonSexViol_i + \beta_3 SexViol_i + \alpha X_i + \varepsilon_{idt} \qquad (2)$$

Where W_{idt} is the dummy variable taking value of one if the individual responded affirmatively to the question of whether (s)he would have been willing to pay a higher price in exchange for a completely safe trip; $\log(price)_i$ is the natural log of the stated price in such question (randomly assigned to individual *i*); $NonSexViol_i$ is the dummy variable taking value of one if the individual reported having experienced any kind of non-sexual aggression; $SexViol_i$ is the dummy taking value of one if the individual reported having experienced any kind of sexual violence; and all other variables and subscripts are defined as before.

Results are reported in Table 8. The first three and last three columns present the results for the pooled women's and men's sample, respectively. Columns 1 and 4 show the OLS estimation results. Thus, with this specification, we recover the demand for safe transportation for both men and women (given that the price was randomly allocated). The other columns show the logit estimates that are employed to calculate the WTP to avoid sexual and non-sexual violence. Columns 2 and 5 show the results without adding controls, while columns 3 and 6 add the controls employed in the main estimations.

Throughout the different specifications, price has a strong predictive power on the likelihood that individuals will respond affirmatively to the question asked. In line with what should be expected, individuals are considerably less likely to declare being willing to pay for a safe trip when the price of said trip is higher. The point estimates are larger in magnitude for the women's sample, suggesting that violence does impose larger costs on them. Columns 1 and 4 show that a 1% price increase

conveys a 0.48 and 0.36 percentage point decrease in the likelihood of affirmatively responding to the WTP question.

The WTP indicators reported result from the ratio of the experienced violence coefficient and the log price coefficient (i.e. $-\beta_2/\beta_1$ for non-sexual violence and $-\beta_3/\beta_1$ for sexual violence). This indicator follows the standard literature practice and is equivalent to the percentage difference in price that a respondent would be willing to pay if she experienced violence with respect to one that did not. Therefore, we use the indicator as a proxy of the cost that violence encompasses. In the case of women, sexual violence conveys a high cost since those that experienced sexual violence are willing to pay 23.6% more. In contrast, the WTP to avoid non-sexual violence is small (4%) and non-significant for women. Interestingly, men also display a large WTP to avoid non-sexual violence (17%), being the estimate significant at a 10% level. Their WTP to avoid sexual violence is also high (25%), but not accurately estimated nor non-significant.

Pairing the WTP results with the impacts shown in section V that suggest a reduction in sexual violence for women equal to 2.9 percentage points and an increase in non-sexual aggression against men equal to 15.3 percentage points, implies that it is not clear that the program is welfare enhancing. Our results show that the program achieves its main objective of reducing sexual harassment against women and that women are willing to pay a high price to avoid sexual violence, making it welfare enhancing for them. However, the program also negatively impacts men at a non-negligible cost, making it pervasive for this group. Thus, we cannot conclude that the benefits of the program outweigh the costs associated with the undesired effects.

Table 8

Willingness to pay Dependent variable: Dummy=1 if willing to pay a higher price						
	Women			Men		
	(1)	(2)	(3)	(4)	(5)	(6)
log(price)	-0.487***	-2.183***	-2.011***	-0.359***	-1.492***	-1.547***
	[0.0431]	[0.172]	[0.187]	[0.0708]	[0.285]	[0.314]
Sexual violence		0.516***	0.470***		0.383	0.339
		[0.141]	[0.145]		[0.363]	[0.397]
Non-sexual violence		0.0912	0.0800		0.257*	0.274*
		[0.0841]	[0.0865]		[0.135]	[0.140]
Willingness-to-pay						
Sexual violence		23.62%***	23.35%***		25.69%	21.90%
		(7.17)	(7.51)		(28.06)	(30.75)
Non-sexual violence		4.18%	3.98%		17.23%*	17.73%*
		(3.82)	(4.77)		(10.30)	(10.10)
Observations	3,419	3,419	3,419	1,497	1,497	1,497
Estimation	LPM	Logit	Logit	LPM	Logit	Logit
Controls						
Date FE	Yes	No	Yes	Yes	No	Yes
Socio-Economic variables	Yes	No	Yes	Yes	No	Yes

Robust standard errors in brackets
Willingness-to-pay SE were calculated using bootstrap with 500 repetitions
* significant at 10%, ** significant at 5%, *** significant at 1%

VIII. Conclusions

The lack of safety in public spaces is likely to impose costs to individuals, and particularly to women, who are considerably more vulnerable to sexual harassment and are more responsive to the fear of victimization. Nonetheless, while having attracted the attention of policy makers and international organizations alike, policy proposals aimed at combatting sexual harassment in public places are rather scarce, and there is a lack of empirical evidence evaluating the existing programs. In particular, there is little discussion regarding the potential undesired effects of this kind of intervention. This paper represents, to our knowledge, the first to empirically assess the impact of women-only public transportation options not only on the prevalence of sexual harassment, but also on non-sexual aggression.

In particular, by exploiting the discontinuity in the hours of operation a program that reserves subway cars exclusively for women in Mexico City, we compare self-reported prevalence of sexual and physical violence experienced by subway users before and after the starts or ends operating within each subway line and day of analysis. We find that, while sexual violence experienced by women is lower when the program is in operation, non-sexual violence experienced by men is higher. Our results then suggest that programs of this kind are likely to have undesired consequences. Unfortunately, our WTP calculations are not precise enough to reach definite conclusions, but they allow us to give evidence of the welfare tradeoff that the program conveys.

References

Agrawal, A. & Sharma, A. (2015). Gender Contests in the Delhi Metro: Implications of Reservation of a Coach for Women. *Indian Journal of Gender Studies*, 22(3), 421-436.

Boarnet, M. G., & Hsu, H. P. (2015). The gender gap in non-work travel: The relative roles of income earning potential and land use. *Journal of Urban Economics*, *86*, 111-127.

Borker, G. (2017). Safety First: Perceived Risk of Street Harassment and Educational Choices of Women. *Working Paper*.

Calonico, S., Cattaneo, M. D., & Titiunik, R. (2014). Robust nonparametric confidence intervals for regression-discontinuity designs. *Econometrica*, 82(6), 2295-2326.

Crane, R. (2007). Is there a quiet revolution in women's travel? Revisiting the gender gap in commuting. *Journal of the American planning association*, *73*(3), 298-316.

Crane, R., & Takahashi, L. (2009). Sex Changes Everything The Recent Narrowing and Widening of Travel Differences by Gender. *Public works management & policy*, *13*(4), 328-337.

Cutler, D. M., & Glaeser, E. L. (1995). *Are ghettos good or bad?* (No. w5163). National Bureau of Economic Research.

Dunckel-Graglia, A. (2013). Women only transportation: How "pink" public transportation changes public perception of women's mobility. *Journal of Public Transportation*, 16(2), 85-105

Ericksen, J. A. (1976). An analysis of the journey to work for women. *Soc. Probs.*, *24*, 428.

Ferraro, K. F. (1995). *Fear of crime: Interpreting victimization risk*. SUNY press.

Garibi, M., Núñez, M. F. Z., & Ríos, E. P. (2010). La discriminacion y violencia contra las mujeres en el transporte publico de la Ciudad de Mexico.

Halpern, D. F., Eliot, L., Bigler, R. S., Fabes, R. A., Hanish, L. D., Hyde, J., ... & Martin, C. L. (2011). The pseudoscience of single-sex schooling. *Science*, *333*(6050), 1706-1707.

Handy, S. (1996, October). Non-work travel of women: patterns, perceptions and preferences. In *e Women's Travel Issues Second National Conference*.

Hanson, S., & Hanson, P. (1981). The impact of married women's employment on household travel patterns: A Swedish example. *Transportation*, *10*(2), 165-183.

Hanson, S., & Johnston, I. (1985). Gender differences in work-trip length: explanations and implications. *Urban geography*, *6*(3), 193-219.

Hjorthol, R. J. (2000). Same city—different options: An analysis of the work trips of married couples in the metropolitan area of Oslo. *Journal of Transport Geography*, *8*(3), 213-220.

Kondylis, Florence, Arianna Legovini and Astrid Zwager (2017). Willing to pay to avoid harassment: Evidence from public transport in Rio de Janeiro. Mimeo.

Johnston-Anumonwo, I. (1997). Race, gender, and constrained work trips in Buffalo, NY, 1990. *The Professional Geographer*, *49*(3), 306-317.

Lynch, G. and Atkins, S. (1988). The influence of personal security fears on women's travel patterns. *Transportation*, 15(3), 257-277.

Martin, C. L., & Fabes, R. A. (2001). The stability and consequences of young children's same-sex peer interactions. *Developmental psychology*, *37*(3), 431.

Mauch, M., & Taylor, B. (1997). Gender, race, and travel behavior: Analysis of household-serving travel and commuting in San Francisco bay area.*Transportation Research Record: Journal of the Transportation Research Board*, (1607), 147-153.

McGuckin, N., & Murakami, E. (1999). Examining trip-chaining behavior: Comparison of travel by men and women. *Transportation Research Record: Journal of the Transportation Research Board*, (1693), 79-85.

O'Flaherty, B., & Sethi, R. (2010). Homicide in black and white. *Journal of Urban Economics*, *68*(3), 215-230.

Sarmiento, S. (2000, January). Household, gender, and travel. In *Women's Travel Issues Second National Conference*.

Steiner, R. (2000). Women's Travel for Shopping in Traditional Neighborhoods: How Does a Woman's Role in the Household Affect Their Activity and Travel for Shopping. In *Proceedings of the Second National Conference on Woman's Travel Issues, 2000*.

Appendix 1. Survey Questionnaire

Pollster Only
Date: _____
Exact Time the Metro Arrived: _____
Metro Line and Station:
_____
Are the women's and men's sections separated? Yes ____ No____
Metro car in which the survey occurred
_____

Pollster Only

Pollster Name: _____
Pollster Code.:_____
Start Time: _____ End Time: _____
Survey Number: _____

"Hello, we are students from the UAM (Autonomous Metropolitan University) and we are administering an anonymous survey about the Metro. May I accompany you while you walk? "

I RESPONDENT'S INFORMATION

1) Age (use numbers)	2) Currently you are:	3) Do you have children?	4) Does your household have a car?
_____	Single ..... 1 Married .....2 Common-Law Marriage……3 Separated or Divorced.....……4 Widow/Widower……...5 Don't Know/Doesn't Answer (DK/DA)……99	Yes……..1 No……2 DK/DA……99 3.1 How Many (number)_____	Yes……..1 No……2 DK/DA……99 4.1 Are you the member of your household who most uses this (these) car(s)? Yes……..1 No……2 DK/DA……99

5) Are you currently a student? Yes ___ No_____ Yes => What level are you studying? NO=> What was the last level you completed?	6) Are you currently working? Yes_____ No _____ 7) What was the main reason for your trip today? *Select more than one if necessary.*	8) How frequently do you make this trip?
Incomplete Elementary school or less ……....1 Elementary…….....2 Secondary….....3 High School/Technical School……...4 University……5 Postgraduate ……...6 DK/DA…….....99	Work……..1 School……2 Shopping……...3 Medical……...4 Social……5 Food……6 Recreation……..7 Returning Home .....8 7.1 Other: List 9	Everyday...…...1 Every weekday.……2 2 to 4 times a week..…...3 Once a week.……5 Less than once a week.....7 Once every 2 weeks……8 Once a month or less……..……9

II TRIP Now we will ask you some questions about your trip TODAY on the [COLOR] line, which you just got off of

9) At what station did you begin your trip on the [COLOR] line? _____ 10) Approximately how long ago did you board the [COLOR] line at [STATION NAME]? _____	11) When you boarded at [STATION NAME] was there a division between men's and women's cars? Yes _____ No _____	12) How was this division established at [STATION NAME]? Police .....1 Physical barrier .....2 Tape on the floor .....3 Sign .....4 DK/DA 99

13) On your trip did you get onto the first train that passed or did you wait for a later train? First train……..1 Second Train……2 Third Train……..3 Other: Which?.....4 13.1 _____ *If the respondent answers that s/he did not get on the first train, ask Question 14. If not, ask Question 15.*	14) Why didn't you get on the first train that passed? *Do not read the options* There were too many people at the station and I could not get to the door ..... 1 The train was full……..2 There was no space to sit on the train……...3 I was waiting for someone……4 The train seemed unsafe……5 It was the women's only car……..6 Other: List.....7 14.1 _____	15) On the train that you boarded at [STATION NAME] were there available seats? Yes……..1 No……2

III Perception of Insecurity during the Trip Now we are going to talk about THE TRIP YOU JUST TOOK ON LINE [COLOR]

16) During your trip on line [COLOR] did you feel unsafe? Yes……..1 No…….2	17) During your trip on line [COLOR] did you observe or experience any of the following situations?									
		Observed		Gender of the Victim		Gender of the Aggressor		Experienced		Gender of the Aggressor	
		Yes	No	Male	Female	Male	Female	Yes	No	Male	Female
	1.Shoving, elbowing										
	2. Physical fighting										
	3. Insults										
	4. Robbery										
	5. Unwanted Sexual looks										
	6. Catcalls										
	7. Photos taken without consent										
	8. Groping without consent										
	9. Brushing against your body										
	10. Any other type of sexual violence _____										

Only ask these questions if the respondent answered that he/she experienced a robbery. If not, go to Question 20. 18) What was stolen? _____ 19) ¿Where? Entering the Metro…..…1 Turnstiles…….2 Platform……3 Train……4 Stairs…..5 Unknown .....6 Other: List…….7 19.1_____	20) If you had not been able to take the Metro today, what method of transportation would you have utilized to arrive at your destination? *Write the codes in the relevant boxes* Metrobús (bus rapid transit) ……1 Large Bus……2 Small Bus……3 Taxi……4 Personal Car/Ride….5 Walking .....6 Light rail……7 Bicycle .....8 Rickshaw .....9 Trolley RTP/Atenea (publically regulated bus)	21) Have you ever taken this trip via this different method of transportation? Yes……..1 No……2 22) ¿How much would this trip have cost using this other method? _____ Mexican Pesos 23) ¿How long would it have taken you to make this trip via this other method? _____ minutes more (23.1) _____ minutes less (23.2)	24) Imagine that there exists a completely safe Metro that spans the entire trip you just took, but that it costs **$6 Mexican pesos.** Given your experience TODAY, would you prefer to pay **$6 Mexican pesos** and take the safe Metro? Yes _____ No _____ 25) Would you be willing to always pay **$8 pesos** for a ticket for a completely safe Metro? Yes _____ No _____

Appendix 2. A simple model

In this appendix, we present a simple model that abstracts from the potential impact of differences in social norms in gender-mixed and single-gender environments and from agglomeration, that delivers the study's findings: an increase in physical violence under single-gender subway cars. In particular, we discuss the implications of Donohue and Levitt (1998)'s model in a context in which there exist observable differences between two groups of individuals, *A* and *B* (men and women).

Consider, then, the following setting:

When an individual boards the subway, (s)he is randomly matched with another individual, with which (s)he engages in a bargaining process in order to obtain a unique and indivisible prize $W>0$ (the private value of a seat, or of personal space in the subway, for example).

Each individual can claim the prize for him/herself, or simply give it away to the other player. If a player gives the seat away, (s)he receives a payoff of zero, while the other player receives *W* if (s)he had claimed the prize.

A conflict occurs when both players claim the prize, in which case it is assigned to the player of the highest "fighting ability". If conflict occurs, each player pays a cost *C*. The loser's payoff is then *–C*, while the winner's payoff is *W-C*.

Each individual is characterized by an observed (θ_i) and unobserved (ε_i) component of his fighting ability, F_i:

$$F_i = \theta_i + \varepsilon_i$$

θ_i is assumed to be observed by both players and to differ in means by individuals' gender, while ε_i is unobserved even by the player him/herself. Further, the θ's are assumed to be normally distributed within each group, independent, and $E(\theta|A) = \theta_A, E(\theta|B) = \theta_B,$. Assume further that $Var(\theta|A) = Var(\theta|B) = \frac{1}{2}\sigma_\theta^2$.

The ε's are assumed to be independently and identically distributed with a type-I extreme value distribution characterized by:

$$Pr[\varepsilon_i \leq \varepsilon] = exp\left[-exp\left(\frac{-\varepsilon}{\sigma_\varepsilon}\right)\right]$$

The distributional assumptions for the unobserved component of the fighting ability allows to derive simple conditions under which conflict will occur, namely:

$$\theta_1 - \theta_2 \geq ln\left(\frac{C}{W}\right)\sigma_\varepsilon$$

and

$$\theta_1 - \theta_2 \leq -ln\left(\frac{C}{W}\right)\sigma_\varepsilon$$

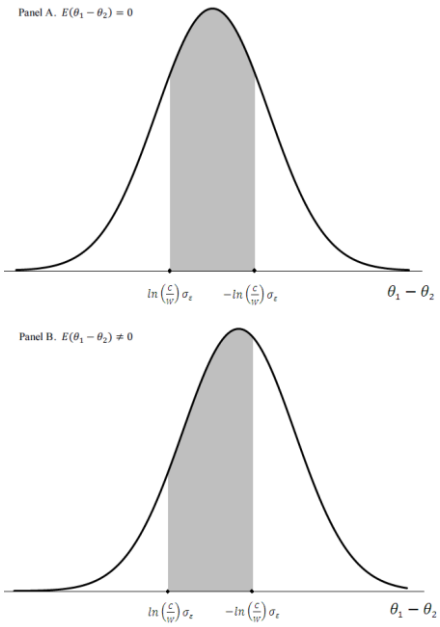
The normality and independence assumption for the distribution of the θ's within each group implies that the distribution of $\theta_1 - \theta_2$ is normal, with variance equal to σ_θ^2. As a result, the probability of a conflict ocurring can be expressed as:

$$P[conflict] = 1 - \Phi\left(-ln\left(\frac{C}{W}\right)\sigma_\varepsilon\right) - \Phi\left(ln\left(\frac{C}{W}\right)\sigma_\varepsilon\right)$$

It is easy to see why, under this setup, conflict will be more likely to occur between matches of individuals of the same group. Under gender separation, $E(\theta_1 - \theta_2) = 0$. However, under integration, for matches between individuals from different groups, $E(\theta_1 - \theta_2) \neq 0$. Figure 1 then illustrates how under integration the prevalence of conflict is minimized.

$$P[conflict|segregation] > P[conflict|integration]$$

Figure A1. Differences in observables and conflict



Appendix 3. Additional Tables

Table A1

Effects on sexual violence

Reduced form regression results: Women					
Dependent Variable: Dummy=1 if sexual violence was experienced					
	Unwanted looks	Sexual comments	Pictures taken	Inappropiate touching	Brushing against body
Pooled	-0.0312**	0.0038	-0.0017	-0.0075**	-0.0006
	[0.0138]	[0.0061]	[0.0022]	[0.0037]	[0.0069]
Morning	-0.0190	0.0025	-0.0014	-0.0102**	0.0069
	[0.0148]	[0.0072]	[0.0016]	[0.0051]	[0.0078]
Afternoon	-0.0565*	0.0045	-0.0040	0.00003	-0.0125
	[0.0322]	[0.0138]	[0.0066]	[0.0012]	[0.0144]
Controls					
Date FE	Yes	Yes	Yes	Yes	Yes
Socio-Economic var	Yes	Yes	Yes	Yes	Yes

Specification using a local polynomial with triangular kernel and 120 minutes bandwidth
Robust standard errors in brackets
* significant at 10%, ** significant at 5%, *** significant at 1%

Table A2

Effects on non-sexual violence

Reduced form regression results				
Dependent Variable: Dummy=1 if physical violence was experienced				
	Pushing / shoving	Fights	Insults	Theft
Women				
Pooled	0.0360	0.0001	0.0190***	-0.0015
	[0.0287]	[0.0054]	[0.0072]	[0.0025]
Morning	0.0047	-0.0037	0.0129	-0.0026
	[0.0388]	[0.0084]	[0.0118]	[0.0028]
Afternoon	-0.0178	0.0001	-0.0029	-0.0027
	[0.0466]	[0.0047]	[0.0070]	[0.0071]
Men				
Pooled	0.1458***	0.0081	0.0210	0.0059
	[0.0431]	[0.0114]	[0.0139]	[0.0086]
Morning	0.2132***	0.0089	0.0244	0.0180
	[0.0509]	[0.0204]	[0.0249]	[0.0169]
Afternoon	0.0439	0.0010	0.0122	-0.0090
	[0.0509]	[0.0016]	[0.0112]	[0.0076]
Controls				
Date FE	Yes	Yes	Yes	Yes
Socio-Economic var	Yes	Yes	Yes	Yes

Specification using a local polynomial with triangular kernel and 120 minutes bandwidth
Robust standard errors in brackets
* significant at 10%, ** significant at 5%, *** significant at 1%