The impact of the MeToo Scandal on women's perceptions of safety

Laila Ait Bihi Ouali[1] and Daniel J. Graham[2]

[1]*Corresponding author: Imperial College London, London, SW7 2AZ, UK. Email:*
`laitbihi@imperial.ac.uk`.
[1,2]*Transport Strategy Centre, Imperial College London*

Abstract

This study assesses whether the negative exogenous informational shock of the MeToo scandal has affected women's perception of safety. The MeToo movement was first reported in the media worldwide in October 2017, and has received enormous press coverage since then. The exogenous and unanticipated nature of the scandal provides a natural experiment that we can use to quantify how wider external information affects 'ordinary' women's perceptions of safety and their willingness to report feelings of dissatisfaction with safety. To do so, we use a case study of women's' perception of safety in 25 metro systems across the world, for which we have large-scale unique customer satisfaction data over the years 2014 to 2018. We use interview dates to determine perceptions pre and post scandal, and we take the MeToo scandal as a 'treatment' affecting women in the sense that they were the primary target of the informational shock. Using pre and post treatment data, for our defined treated (women) and control (men) units we apply a difference-in-differences estimator to identify the impact of the scandal on perceptions of safety. Our results show an 2.5% increase in the probability of women being dissatisfied with security in stations and in metro carriages post-scandal. These results support the conclusion that revealed preferences are only revealed up to a certain extent. A change in the context (*e.g.*, an informational shock), can alter perceptions and in turn, can encourage individuals to disclose lower (or higher) satisfaction levels as perceptions of the norm change.

Keywords: Gender; Perceptions; Safety, Security; Metro; Revealed Preferences; Stated Preferences.

1 Introduction

People's perceptions and preferences depend on the nature and context of the information available to them. Within the abundant literature on individual responses to additional information, many studies explore the importance of inferring from motivated reasoning (*e.g.* Kunda 1990, Bénabou and Tirole 2002, 2011). This concept proposes that even in settings in which individuals are constrained to be accurate, they remain unable to control their own preconceptions when making a decision. This view of belief formation is consistent with the renewal of interest in emotions and their influence on decision-making in psychology and neuroscience (*e.g.*, Dolan and Sharot (2011)). Motivated reasoning is different than cognitive dissonance and confirmation bias, representing situations in which people overweight their prior beliefs when updating their views (Festinger 1957, Tetlock 1983, Rabin and Schrag

1999, Taber and Lodge 2006). Yet, confirmation bias can coexist with motivated reasoning and there is ample experimental evidence of this form (*e.g.* Taber and Lodge 2006, Kahan et al. 2012, Nyhan and Reifler 2010, Nyhan et al. 2013).

Motivated reasoning implies that individuals do not fully control their own preconceptions when making decisions, which creates a tension between accuracy and the perseverance of beliefs. Under the assumption of rationality, the collection and integration of new information should be independent of one's prior judgement (Evans and Over 2013). Depending on the credibility of a piece of evidence, individuals' perceptions should be adjusted upwards or downwards (Anderson 1981). However, individuals are not purely rational.

Our starting premise is that all reasoning is at least in part motivated (Kunda 1987, 1990), and that consequently we expect an informational shock to alter perceptions. Since its unravelling on October 15, 2017, the MeToo movement has incentivised individuals to reveal experiences of harassment, whether they experienced it themselves or have shared someone else's experience(s). Although this event initially started in America, the press coverage around these themes spread worldwide (see Fig. 1) and sparked wider debates around women's security. Since the scandal, individuals have been actively searching this topic on the internet, indicating that they are aware of the movement and its content as well (see Fig. 2). As such, it constitutes an excellent candidate for a quasi-experimental study of the impacts of information on perceptions.
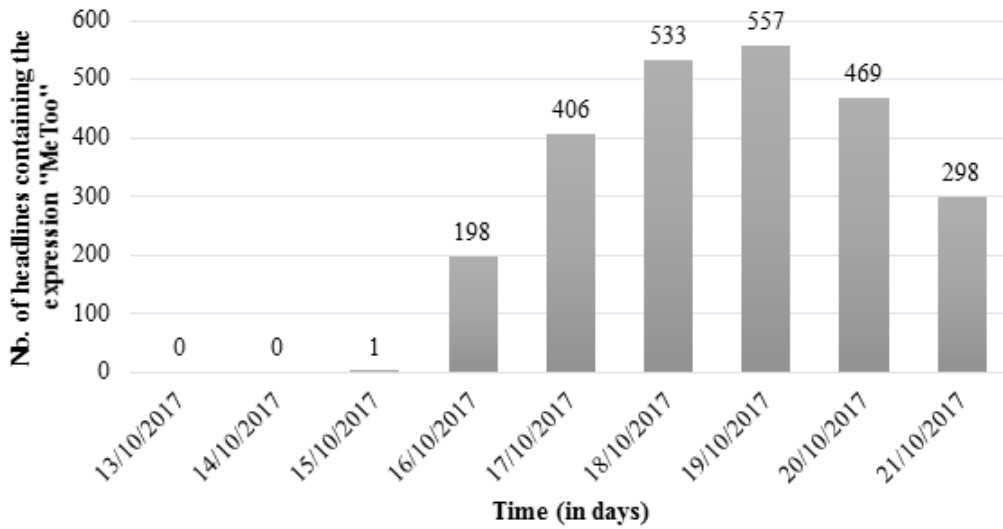

Figure 1: Detailed daily count of headlines containing the keywords "MeToo" (source: Factiva)
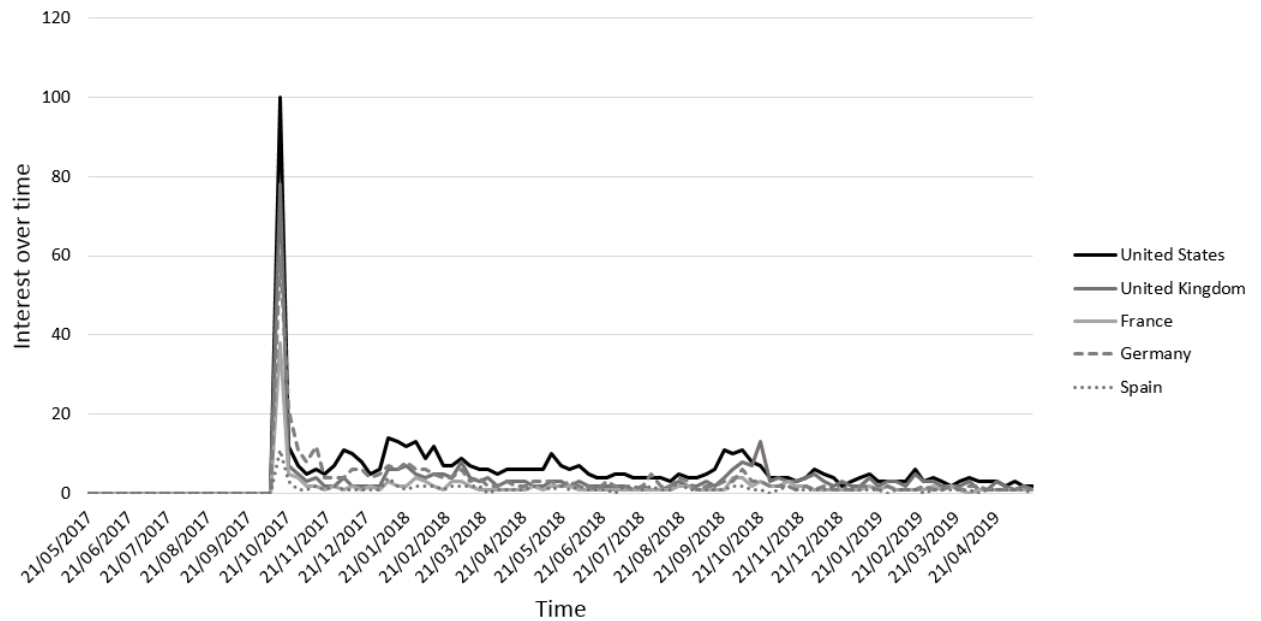
Figure 2: Web search intensity for the keywords "MeToo" (source: Google Trends)

1 We exploit the exogenous and unanticipated nature of the MeToo scandal, and the unique
2 setting surrounding this event (intensive press coverage, worldwide magnitude of the event), to
3 identify changes in women's perceptions and propensity to reveal their preferences. Building
4 upon the principles of motivated reasoning, our analysis tests empirically the extent to which
5 behaviour changes post-scandal. Our specific research questions are as follows. Did the
6 informational shock associated with the MeToo scandal change women's propensity to state
7 their level of contentment with safety? Do negative informational shocks leads to a significant
8 update in the stated preferences?

9 To answer these questions we conduct a case study of women's' perception of safety in 25 urban
10 metro systems around the world, for which we have large-scale unique customer satisfaction
11 data over the period 2014 to 2018. While the MeToo movement had no specific association with
12 public transport, women's safety on public transport has been of long standing concern, and
13 in fact as Figure 3 shows, articles with headlines referring to MeToo and transport increased
14 massively after the MeToo scandal. Thus, we believe that our unique data provide a route to
15 examine the effects of MeToo on the perceptions of 'ordinary' women that use public transport
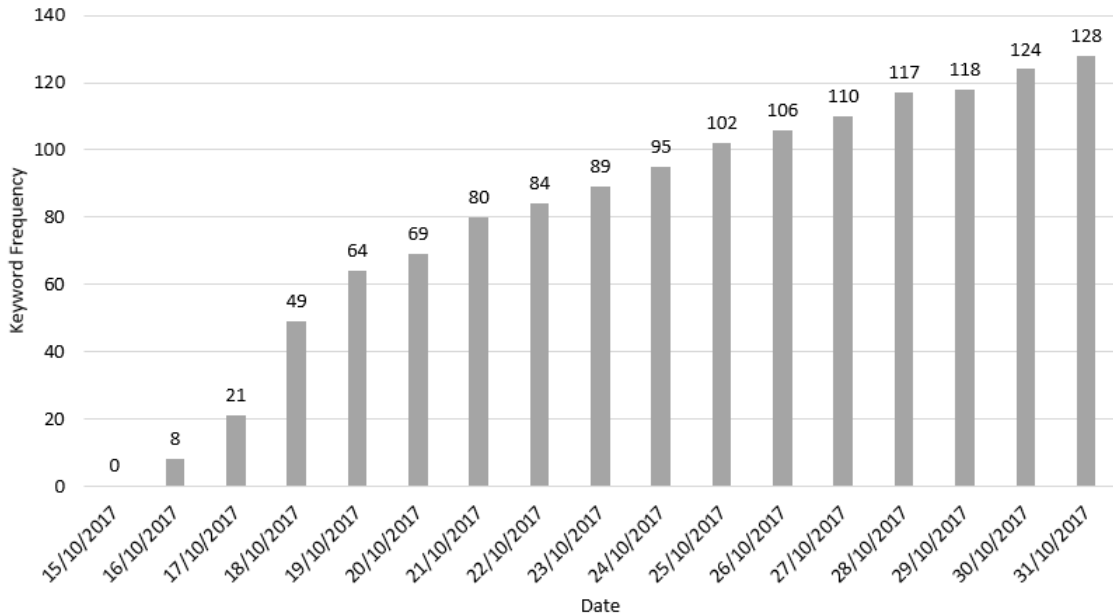16 systems to go about their daily business.

Figure 3: Cumulated number of headlines with keywords related to Me Too and transport (metro, transport, public transport) – Source: Factiva

1 The paper makes two main contributions. The first is empirical. To our knowledge we are
2 the first statistical study to explore the impact of the MeToo scandal on the perception of
3 safety in public transport worldwide. Our data on the stated preferences of over 137,500
4 individuals from 25 cities across the world provides an excellent opportunity to explore this
5 issue in a detailed and comprehensive fashion. Our second contribution lies in the nature
6 of our statistical approach we use to identify the impact of MeToo on women's perceptions
7 of safety. We define the MeToo scandal as 'treatment' creating pre- and post- treatment
8 periods, and we define women as treated 'units' and men as controls. However, we also
9 acknowledge that men could be affected by this informational shock and incentivised into
10 revealing different preferences. To deal with this seeming inconsistency, we use a difference-
11 in-differences methodology which provides us with clear-cut causal estimates and eliminates
12 trend biases, while also controlling for men's potential reactions to the informational shock.
13 Our DiD analysis can also explore the existence and extent of differentiated responses to test
14 how informational shocks affect subgroups based on socio-demographics.

15 Using pre and post scandal (treatment) data, for our defined treated (women) and control
16 (men) units, we apply a difference-in-differences (DiD) estimator to identify the impact of the
17 scandal on perceptions of safety. Our results show evidence that the MeToo scandal did lead
18 to an increase in the propensity of women to reveal and state concerns over safety. We find this
19 effect essentially for respondents in European countries for which we find an average increase
20 of 2.5% in women's propensity to agree that metro stations are not secure, following the
21 MeToo scandal. Our DiD results are stable across specifications and to location controls and
22 time fixed effects. Testing for heterogeneity in responses show that there are no heterogeneous
23 effects by age levels, and that women's reaction is of a larger magnitude if they are daily metro
24 users.

25 Our results support the conclusion that revealed preferences are only revealed up to a certain

4

extent, and that informational shocks encourage individuals to disclose different satisfaction levels as their perception of the norm change. The paper is structured as follows. Section 3 presents the data and Section 4 the empirical model. Section 5 presents and interpret results along with robustness checks and sensitivity analyses. Section 6 concludes.

2 The MeToo Scandal

The MeToo movement was first reported in the media worldwide in October 2017, and has been subject to a substantial press coverage since, as presented in Figures 2 and 1 that indicate respectively an increased web search intensity and the number of headlines. While the MeToo movement had no specific association with public transport, women's safety on public transport has been of long standing concern. One might object that the MeToo movement was not associated with public transport in articles, yet Figure 3 shows that articles with headlines referring to MeToo and transport started and increased massively after the MeToo scandal[1].

In this study, the MeToo scandal is analysed from a perspective that is moved away from testimonies within the Hollywood industry and is closer to the public debate. Our rationale revolves around Europe as public transport is widely used and spread in these countries, and metros are used by general population and across all walks of life. Therefore, it is relevant to focus our study on public transport patronage in Europe as it differs from that in Asia and the Americas, with customer data in European transport systems being more representative of the population. Thus, we believe that this unique event provides a route to examine the perceptions of women that use public transport systems to go about their daily business.

To the best of our knowledge, there has been no quantitative analyses of the impact of the MeToo movement. Analyses led so far are quite scarce, of a qualitative nature and do not analyse individual responses to the scandal at a large scale. More precisely, one study present the interest of an application aimed at helping harassed women (Sheikh and Fayyaz 2019) in the context of MeToo but does not analyse the phenomenon in itself. Another study uses the MeToo context to analyse qualitatively how anger is mediated after the scandal but uses the MeToo scandal from a Hollywood viewpoint to analyse individual responses (Orgad and Gill 2019). Gu et al. (2019) refers to the MeToo scandal to mention gender differences and discrimination but does not study the scandal *per se*.

Prior research has tested the impact of informational shocks on individuals (Sides 2011, Cruces et al. 2013, Kuziemko et al. 2015, De Neve et al. 2017). However, individual responses to informational shocks are often studied in the field of public economics and more often it is its impact on voting behaviours that is explored (Ait Bihi Ouali 2019). According to Alesina and Giuliano (2011) and Margalit (2013), negative experiences can change individuals' preferences and make them less optimistic or more risk averse. Fiscal scandals are considered a negative experience, triggering a revision of beliefs.

Our previous research shows that women have different priors than men on the perception of

---

[1]Shortly after the scandal first broke in the media, a survey suggested that individuals are more willing to speak out against harassment post-MeToo scandal, which already hints at the fact that this informational shock affects individual perceptions, see: $https://www.fawcettsociety.org.uk/metoo-one-year$.

safety in public transport. Ait Bihi Ouali et al. (2019) find that for both buses and metros worldwide, women are 6% to 10% more likely to be dissatisfied with security in public transport than men. One's prior beliefs and attitudes can have an important bearing on the evaluation of new information. Given that women already have priors that are more negative than men's, we would expect an informational shock to affect women differently because motivated reasoning leads people to become more polarized.

3 Data

3.1 Customer Satisfaction Surveys

The data used in this study have been collected annually over the period 2014 to 2018 via responses to Customer Satisfaction Surveys (CSSs) from users of urban metros. Data collection is facilitated through collaboration between the Transport Strategy Centre (TSC) at Imperial College London and several major public transport operators, organised in the form of two consortia of urban metros, CoMET (Community of Metros) and Nova.

The CoMET group is a consortium of some of the world's largest urban metros including: Berliner Verkehrsbetriebe (BVG, Berlin), Delhi Metro Rail Corporation (DMRC), Mass Transit Railway (MTR, Hong Kong), the Underground (London), Sistema de Transporte Colectivo (STC, Mexico City), Metro de Madrid, Moscow Metro, New York City Transit (NYCT, New York), the Régie Autonome des Transports Parisiens (RATP, Paris) that include both the Metro and the Réseau Express Régional (RER), Metro de Santiago and the Singapore Mass Rapid Transit (SMRT, Singapore). The Nova group is a consortium of small to medium sized metros including: Buenos Aires Metrovias, Transports Metropolitans de Barcelona (TMB, Barcelona), Société des Transports Intercommunaux de Bruxelles (STIB, Brussels), Bangkok Expressway and Metro Public Company (BEM, Bangkok), Docklands Light Railway (DLR, London), Istanbul Ulasim, RapidKL / Prasarana (Kuala Lumpur), Metropolitano de Lisboa, Société de Transport de Montréal (STM, Montréal), Newcastle Nexus, Metro Rio (Rio de Janeiro), Toronto Transit Commission (TTC, Toronto) and Vancouver SkyTrain (Vancouver, Canada).

The first part of the CSS questionnaire contains statements relating to eight customer service areas defined under European Norm 13816: availability, time, information, comfort, security[2], customer care, accessibility and environment (see EN13816 2002). In addition, there is one general question on overall satisfaction. The questionnaires are produced and disseminated via an online survey building and hosting tool. Where necessary, translations of the survey are provided by operators into their home languages. For example, for multi-language cities such as Brussels, Belgium, there are separate surveys produced for each major language. Participating operators posted the link(s) to their own survey(s) on their home page for the same 4-week period each year, or via their social media pages or email bulletins. The important argument in favour of comparability is contained in the consistency of this method: face-to-

---

[2]We acknowledge that from a practitioners' viewpoint, definitions of safety and security differ where safety refers to infrastructure failure and security defines unlawful acts interfering with individuals. However, to align our interpretations with the academic literature in urban and transport economics, we will refer to the *perception of safety*.

face or phone interviews are never led in order to avoid interaction with staff that could bias the results of the respondents.

Respondents are asked to provide their opinions on normal service operations. In the first section, answers are given on a scale from "1 - Agree Strongly" to "5 – Disagree Strongly". The second section of the survey asks respondents to select, in order of preference, the three customer service areas that are most important to them. Finally there are four demographic questions to understand the sample frame.

Data available for estimation encompass information yearly for a group of respondents who are customers of urban metros. Metro data records a total of 137,513 observations for 25 cities over the period going from 2014 to 2018. Descriptive statistics in Table A.1 indicate that the distribution of genders is rather equal for both datasets for all years. We observe that individuals in this survey are often daily commuters who are travelling for work. Note that due to the need to respect confidentiality it is not possible to reveal the name of particular transport systems in our analyses. We use the 16 October 2017 as the cutoff date after which the MeToo scandal spread in the media. We have enough data before and after the cutoff date for each subgroup as presented in Table A.2 in Appendix. This indicates that the timing of our data relates to that of the MeToo scandal.

4   The Model

This study uses a difference-in-differences (DiD) empirical strategy to infer causality post scandal. A regression discontinuity design cannot be used because the structure of the data does not permit it: were high-frequency daily individual level data around the cutoff date available, it would have been possible to explore regression discontinuity effects. A DiD estimation strategy is appropriate to test for responses in our setting since the MeToo movement could also potentially affect men's perceptions, which are controlled for in the DiD specification via the common trend. The specification used is as follows:

$$y_{i,t} = \alpha + X_i\beta + \delta Post_t + \gamma Treated_i + \nu(Post_t \times Treated_i) + \epsilon_{i,t} \tag{1}$$

Where $y_{i,t}$ represents the dependent variable, X_i is the vector of sociodemographic characteristics, $Post_t$=1 if the respondent answers after the scandal and 0 otherwise, $Treated_i$=1 if the respondent is a female and 0 otherwise, and $\epsilon_{i,t}$ is the error term. The parameters α, β, δ, γ, and ν are to be estimated. The main coefficient of interest for our study is ν which identifies the causal effect of the MeToo scandal. A priori we expect $\nu < 0$: a negative informational shock lowers the idea of what a standard norm is and decreases women's incentive to express satisfaction with either security or overall satisfaction.

The regressions reported in this paper make use of different dependent variables all of which are derived from statements on security. Statements are suggested to respondents who rank the extent to which they agree with them and this is used as a means to express their preferences. As such, the explanatory variables in our study are discrete, representing individual opinions on safety and satisfaction in public transport. As such, they have non-normal error

distributions and so we use Generalised Linear Models (GLMs) for analysis, and specifically ordered probit specifications for ordinal variables.

While the dependent variables are ordinal, they are not continuous in the sense that the metric used to code the dependent variables encompass different satisfaction levels. Satisfaction levels are represented on a 5-point scale and assigns numbers $\{1, ..., 5\}$ to the categories "Strongly Disagree",..., "Strongly Agree". The metric relating numbers (from 1 to 5) is linear whereas the metric underlying the satisfaction scale is not. For instance, the difference between 1 and 3 ("Strongly Disagree" to "Neither Agree nor Disagree") is likely to be quite different from the difference between 2 and 4 ("Disagree" to "Agree").

The main GLM we consider is a standard response model in which the cumulative probabilities of the discrete outcome are related to an index of explanatory variables. Let y_i be the observed ordinal variable, then we model $\Pr[y_i \le j | x] = \phi(\alpha_j - x_i'\beta)$ with $j = \{1, ..., 5\}$, where α_j and β are model parameters to be estimated and ϕ is the standard normal cumulative distribution function. We assume there is a latent continuous metric behind the observed ordinal responses.

The observed explanatory variable y_i can take values from 1 to 5 such that

$$y_i = j \iff \alpha_{j-1} < y_i^* < \alpha_j$$

with $j = \{1, ..., 5\}$, and where α designates the cutpoints estimated by the data. Cutpoints help in matching the probabilities associated with each discrete outcome. However, the metric of the observed variable y_i is linear but the satisfaction scale is not. Therefore, we assume there is a latent continuous metric behind the observed ordinal responses.

The latent continuous variable y^* represents the satisfaction scale. It is a linear combination of predictors and an error term and can be written

$$y_i^* = x_i\beta + \sum_{k=1}^{N} \delta_k \ c_k + \sum_{m=2009}^{M} \gamma_m \ t_m + u_i \tag{2}$$

where y_i^* is the dependent variable (i.e. a statement on safety or satisfaction in public transport), u_i is the error term, x_i is a vector of observable individual characteristics, $\sum_{k=1}^{N} c_k$ is a set of city specific dummy variables, and $\sum_{m=2009}^{M} t_m$ is a set of time dummy variables for each year.

Following this model, the probability of observing outcome j corresponds to the probability that the estimated linear function, plus random error, is within the range of the cutpoints estimated for the outcome

$$
\begin{aligned}
Pr(y_i = j) &= Pr(\alpha_{j-1} < y_i^* \le \alpha_j) \\
&= Pr(\alpha_{j-1} < x_i\beta + u_i \le \alpha_j) \\
&= Pr(\alpha_{j-1} - x_i\beta < u_i \le \alpha_j - x_i\beta) \\
&= \phi_{ij} - \phi_{ij-1}
\end{aligned} \tag{3}
$$

where ϕ represents the cumulative distribution function in the standardised normal distribution. The estimation of the regression coefficients in vector β is achieved by maximum likelihood.

5   Results

5.1   Baseline Results

Table 1 tests for a preliminary effect in the form of a variation of women's response post MeToo compared to men's. The difference-in-differences indicator amounts to -0.07, which indicates a decrease in the perception of safety post scandal for the treated group (women). This result is consistent with a decrease in women's customer satisfaction after the scandal broke out.

	Pre "Me Too"	Post "Me Too"	Post-Pre	Diff.-in-Diff.
Control Group	3.250842	3.237402	-0.01344	**-0.06777**
Treated Group	2.921891	2.840681	-0.08121	

Table 1: Descriptive Statistics: Preliminary estimations of MeToo on the Perception of Security – Source: CSS data (Comet/Nova), own calculations

Table 2 presents ordered probit estimates using a difference-in-differences methodology. For all regressions in Columns (1) to (5), the coefficient on the gender dummy variable "women" is negative. This corroborates the previous results in Ait Bihi Ouali et al. (2019) which indicate that men and women have intrinsically different priors on their perception of security, with women stating lower security levels compared to men. Table 2 records a drop in stated satisfaction post MeToo for security items and satisfaction questions for European cities. Estimates suggest that after the MeToo hashtag broke into the media, women revealed lower satisfaction levels with security in trains. The MeToo scandal appears to have had an effect observable especially for Europe.

Apart from the likelihood to recommend the service to someone else, the scandal decreased the level of satisfaction overall as well as security for women post MeToo. Estimates indicate that the decrease in stated satisfaction is larger for security in trains (Column 1) than for security in stations (Column 2). This is in line with women's greater perception of the risk in more confined spaces (Jorgensen et al. 2013, Steinmetz and Austin 2014): trains are more confined than stations, and harassment in crowded trains has been widely covered in the media alongside the MeToo scandal. Yet, the magnitude of the response to the MeToo scandal has the largest magnitude for the question on overall satisfaction (Column 3). Ordered probit estimates' value cannot be interpreted as such. Therefore, in order to quantify the change in the probability to pick an answer towards a statement, is necessary to compute marginal effects.

VARIABLES	Trains secure	Stations secure	Overall Satisf.	Likelihood Reco.
	(1)	(2)	(3)	(4)
Female (Yes=1)	-0.213***	-0.226***	-0.0244**	0.00354
	(0.0109)	(0.0106)	(0.0120)	(0.0117)
Post Me Too (Yes=1)	0.113***	0.0690**	0.231***	0.162***
	(0.0268)	(0.0269)	(0.0388)	(0.0286)
Post Me Too & Female (Yes=1)	**-0.0694***	**-0.0446**	**-0.0744***	**-0.0351**
	(0.0215)	**(0.0213)**	**(0.0277)**	**(0.0214)**
City Controls	Yes	Yes	Yes	Yes
Sociodemographic Controls	Yes	Yes	Yes	Yes
Observations	93,240	93,276	93,463	92,193

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table 2: Difference-in-diff. estimates on security and satisfaction statements in Europe – Source: CSS data (Comet/Nova)

1 Marginal effects related to ordered probit specifications are shown in Figure 4. They are
2 computed respectively for statements related to safety in trains and stations, as well as for
3 overall satisfaction. The largest recorded increase is recorded for the item "strongly disagree".
4 We record a much more limited response with respect to neutral statements, with most of the
5 response being concentrated in the decrease in the "Agree" section. For women post-scandal,
6 we record a 2.5% increase in the probability to say that stations are not secure and a 2.5%
7 increase in the probability to state that trains are not secure for them. We also record a 2.1%
8 increase in individuals' probability to state they are dissatisfied overall. Results are robust to
9 the use of an ordered logit specification as well: the corresponding estimates are presented in
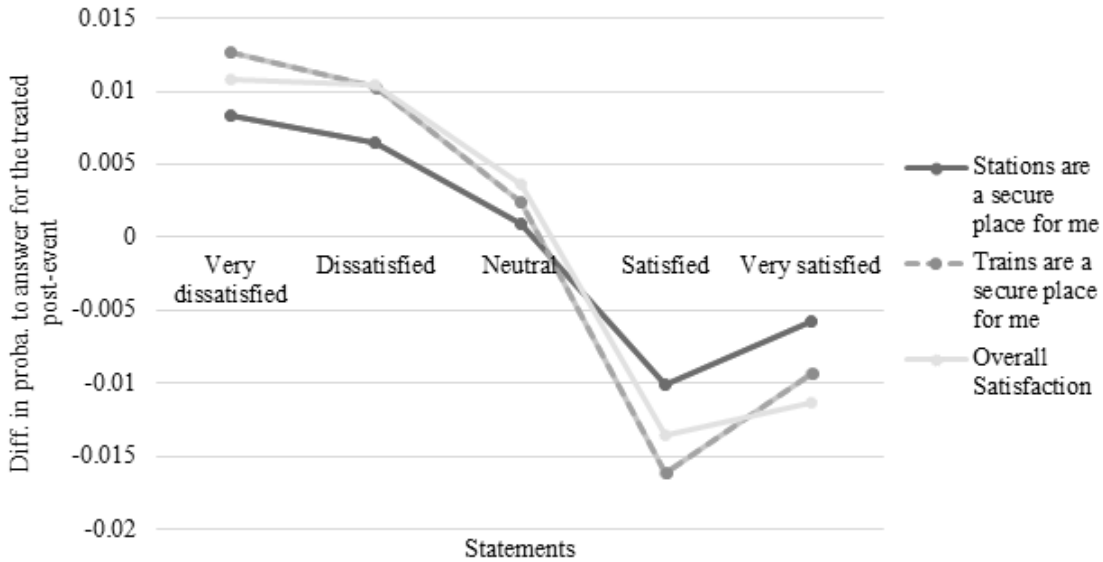10 Table B.3 in Appendix.

Figure 4: Ordered probit estimates' marginal effects – Source: CSS data (Comet/Nova)

10

5.2 Sensitivity Analysis

So far, our results have been estimated conditional on time and city fixed effects in order to assess the main impact of the MeToo paper on the perceptions of safety. However, the response recorded represents an average effect across individuals. While these first ordered probit estimates are helpful, it is interesting to lead complementary analyses to test for heterogeneous effects with respect to several dimensions: (i) sociodemographic variables (age, travel frequency, travel motive), (ii) number of acts of violence, and (iii) press coverage intensity. These regressions follow an ordered probit specification in a triple differences framework.

All specifications include sociodemographic, time and city controls that could be relevant for women' perception of safety. The intensity of the response to the MeToo scandal on the stated perception of safety in public transport, can be determined by the gender, exposure to the scandal and yet another variable. A triple difference strategy may reduce the bias in the estimate of the effect of the scandal. The equations corresponding to the specifications are presented in all the following categories.

5.2.1 Heterogeneity of responses with respect to sociodemographic variables

On the basis of equation 1, the triple differences specification tests whether there is a change in the magnitude of individual responses with respect a sociodemographic variable that is successively age, travel frequency, travel purpose. The specification is as follows:

$$y_{i,t} = \alpha + X_i\beta + \delta Post_t + \gamma Treated_i + \nu(Post_t \times Treated_i * SocioDem_i) + \epsilon_{i,t} \qquad (4)$$

Where $y_{i,t}$ represents the dependent variable, X_i is the vector of sociodemographic characteristics, $Post_t$=1 if the respondent answers after the scandal and 0 otherwise, $Treated_i$=1 if the respondent is a female and 0 otherwise, and $\epsilon_{i,t}$ is the error term. $SocioDem_i$ represents the value of a given sociodemographic variable that are successively age, travel frequency and travel motive depending on the specification. The parameters α, β, δ, γ, and ν are to be estimated. The main coefficient of interest for our study is ν which identifies the causal effect of the MeToo scandal for women by sociodemographic variable.

VARIABLES	Stations are secure for me (1)	Trains are secure for me (2)	Overall Satisfaction (3)
Reference category: 18-29y.o.			
Post & Female & $\leq$ 18y.o. (Yes=1)	0.0738	0.0617	0.0858
	(0.123)	(0.120)	(0.119)
Post & Female & 30-39 (Yes=1)	0.0121	0.0260	0.0433
	(0.0545)	(0.0627)	(0.0644)
Post & Female & 40-49 (Yes=1)	0.00283	0.0201	0.122*
	(0.0621)	(0.0653)	(0.0635)
Post & Female & 50-65 (Yes=1)	0.0147	0.0455	0.146**
	(0.0491)	(0.0515)	(0.0626)
Post & Female & $\geq$ 65 (Yes=1)	-0.00607	0.112*	0.244***
	(0.0661)	(0.0626)	(0.0771)
Observations	93,276	93,240	93,463

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table 3: Age heterogeneity effects across sociodemographic characteristics on secure stations – Source: CSS data (Comet/Nova)

Table 3 shows results testing for heterogeneity with respect to age. The age category of reference is the group of 18 to 29 year old respondents. Results indicate that people who are over 65 respond significantly less compared to the reference group (18-29). However, for overall satisfaction, we notice no significant difference between individuals who are less than 18 to 39.

Table C.5 in Appendix presents heterogeneity estimates with respect to travel frequency, with individuals who travel often as the reference category. Estimates indicate that rare users tend to react significantly less than individuals who travel often. The 'Other' category is one of the smallest category and amounts to 2% of the total group. More precisely, it groups essentially people who travel very occasionally, for instance in the context of an incident where travelling by metro constitutes a second best option (*e.g.*, after an individual's car broke).

When it comes to heterogeneity of responses to the MeToo scandal with respect to the main travel purpose, Table C.5 indicate that individuals who travel mainly to go to school or university react significantly less than those who mainly travel for work. In addition, individuals who travel mainly for medical reasons react significantly less compared to the reference category, probably because they use it less often overall.

5.2.2 Heterogeneity with respect to acts of violence

We draw on Key Performance Indicators data (KPI) that covers extensively 25 metro systems for 2014-2018 to test for heterogenous responses with respect to the number of acts of violence recorded in a given network. Using KPI data, we control whether we can observe heterogeneous responses with respect to the number of acts of violence present in the metro network.

Following a triple differences specification as presented in equation 6 as follows

$$y_{i,t} = \alpha + X_i\beta + \delta Post_t + \gamma Treated_i + \nu(Post_t \times Treated_i * ActsViolence_i) + \epsilon_{i,t} \quad (5)$$

where $ActsViolence_i$ is the number of acts of violence committed in a given metro network. It is treated as a continuous numerical variable in the regressions, and other variables follow the same definition as in equation 6.

Estimates from Table 4 suggest that there is no unobserved heterogeneity with respect to the number of acts of violence. This result corroborates the fact that the change in responses appears to be more likely to be caused by the the informational shock irrespectively of how intrinsically unsafe is a given metro network.

VARIABLES	(1) Stations are secure for me	(2) Trains are secure for me
Female (Yes=1)	-0.217***	-0.200***
	(0.0140)	(0.0138)
Post MeToo & Female (Yes=1)	-0.0554*	-0.0644**
	(0.0311)	(0.0296)
Nb. Acts Violence	-3.14e-05	9.59e-06
	(4.20e-05)	(3.94e-05)
Post & Nb. Acts Violence	-5.49e-05**	-6.24e-05***
	(2.39e-05)	(2.28e-05)
Female & Nb. Acts Violence	-1.66e-05	-2.02e-05*
	(1.08e-05)	(1.18e-05)
Post & Female & Nb Acts Violence	**3.62e-05**	**2.28e-05**
	(2.34e-05)	**(2.27e-05)**
Observations	82,125	82,087

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table 4: Sensitivity Analysis – Ordered probit estimates using dependent variables related to security – Source: CSS data (Comet/Nova)

5.2.3 Heterogeneity with respect to press coverage intensity by country

To test whether the channel triggering the change in individual responses is an informational one, we lead complementary analyses. The media coverage of the MeToo scandal was intensive and very negative, but as presented in Figures 2, 1 and 3, its intensity differed by country. In addition, our specification controls for city fixed effects which eliminates the intrinsic heterogeneity of responses at the city or country level. Following a triple differences specification as presented in equation 6, this specification is as follows

$$y_{i,t} = \alpha + X_i\beta + \delta Post_t + \gamma Treated_i + \nu(Post_t \times Treated_i * PressCoverageIntensity_i) + \epsilon_{i,t} \quad (6)$$

where $PressCoverageIntensity_i$ is the number of headlines covering the MeToo scandal in a given country.

Press coverage is defined as the number of MeToo headlines over the month when the scandal broke out in the media which covers the time span going from October 21st, 2017 to November 21st, 2017. Press coverage intensity differs by country, which is why it is represented as a continuous variable. Results in Table 5 show that the more intensive the press coverage, the

more women tend to state lower safety levels in metros. The triple interaction term captures the impact of the press coverage intensity for women post MeToo scandal.

Columns (1) and (3) in Table 5 indicates the impact of media coverage and suggests that the greater the media coverage, the larger is the decrease in the stated perception of safety post-scandal. Therefore, these complementary estimations reveal that more media coverage by country increase the magnitude of the response. Interestingly, the response affects significantly individuals' overall satisfaction levels.

VARIABLES	Stations are secure for me (1)	Trains are secure for me (2)	Overall Satisfaction (3)
Female (Yes=1)	-0.179***	-0.131***	-0.119**
	(0.0412)	(0.0421)	(0.0473)
Female (Yes=1) & Press coverage	-6.66e-05	-0.000187	0.000238
	(0.000125)	(0.000123)	(0.000151)
Post & Female (Yes=1)	0.0646	0.00958	0.133*
	(0.0595)	(0.0567)	(0.0714)
Post & Female (Yes=1) & Press coverage	**-0.000327****	**-0.000178**	**-0.000691***
	(0.000156)	**(0.000144)**	**(0.000187)**
Observations	26,383	26,374	26,418

Table 5: Sensitivity Analysis – Ordered probit estimates using dependent variables related to press coverage – Source: CSS data (Comet/Nova)

5.3 Common Trend Test

The difference-in-differences strategy consistently captures a significant decrease in the perception of safety in metros after the MeToo scandal. This strategy requires a common trend assumption between the treated and control groups. A standard placebo test would estimate the DiD using cutoff dates different from October 17th, which would enable us to test for false positives. However, use of such a sensitivity test is not ideal as it allows us to test only for one date at a time. Therefore, instead we use a Lags and Leads test as a placebo test. The corresponding specification is as follows

$$y_{i,t} = \alpha + X_i\beta + \delta Treated_i + \sum_k \nu_k Treated_i * LeadsLags_k + \epsilon_{i,t} \qquad (7)$$

Where the $LeadsLags_k$ represent a given month. We expect that there is no significant change in statements among women before the scandal and significant variations in statements among women post-scandal, which can be translated as follows:

1. $\sum_k \nu_k$ jointly insignificant if $k < 0$

2. $\sum_k \nu_k$ jointly significant if $k \geq 0$

We compute F-tests of the joint significance are available in Table 6 below.

	Trains are a secure place for me		Stations are a secure place for me	
Lags	*chi2(7)*	7.04	*chi2(7)*	13.73
	Prob >chi2	0.4246	*Prob >chi2*	0.0562
Leads	*chi2(3)*	19.97	*chi2(3)*	12.21
	Prob >chi2	0.0002	*Prob >chi2*	0.006

Table 6: Heterogeneity of responses with respect to the priority given the priority given to security for dependent variables "Stations are a secure place for me" and "Trains are a secure place for me" – Source: CSS data (Comet/Nova)

1 Table 6 above indicates that for both items related to security, the lags are jointly insignificant
2 and the leads are jointly significant. This is in line with the fact that the reaction of individuals
3 is localised essentially after the MeToo Movement.

4 ## 5.4 Falsification Tests

5 Falsification tests use the same specification as in equation 1. Dependent variables are state-
6 ments that cover the political action on matters different from security. Table 7 presents
7 falsification tests' results for the main outcomes. Results suggest that the opinion/preferences
8 of individuals on other dimensions do not significantly change after the MeToo movement
9 broke out. This result highlights the security dimension as the main channel for the update
10 of beliefs post information shock.

	Available Convenient Network (1)	Easy to get Info on Disruption (2)	Metro Reliability (3)	Likelihood to Reco. to Relative(s) (4)
Female (Yes=1)	0.0293***	-0.0242***	-0.0949***	0.00387
	(0.00926)	(0.00886)	(0.0124)	(0.0115)
Post Me Too (Yes=1)	0.145***	0.210***	0.122*	0.126**
	(0.0415)	(0.0587)	(0.073)	(0.0499)
Post Me Too & Female (Yes=1)	**-0.0164**	**-0.0336**	**-0.0205**	**-0.0356**
	(0.0246)	**(0.0208)**	**(0.0223)**	**(0.0218)**
City Controls	Yes	Yes	Yes	Yes
Sociodemographic Controls	Yes	Yes	Yes	Yes
Observations	93,325	91,406	93,340	92,193

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table 7: Falsification Test – Ordered probit estimates using dependent variables unrelated to security – Source: CSS data (Comet/Nova)

11 # 6 Conclusion

12 This study explores the impact of the MeToo scandal as an exogenous informational shock
13 on women's perceptions of safety using using large-scale unique customer satisfaction surveys
14 about security in urban metros worldwide. The MeToo movement provides a unique setting
15 in which intensive press coverage can potentially trigger a revised understanding of women's

security and satisfaction. Following MeToo, we find a change in stated preferences for individuals in Europe in the form of a decrease in stated satisfaction with safety in metros. We record a 2.5% increase in the probability to say that stations are not secure and a 2.5% increase in the probability to state that trains are not secure for them. We also record a 2.1% increase in individuals' probability to state they are dissatisfied overall with the metro system. Results are stable across specifications, to location controls and time fixed effects. We find that effects are heterogeneous with respect to age, travel purpose and travel frequency. The revised reaction of women is of a larger magnitude if they use the service daily.

Additional tests quantify whether the magnitude of responses changes across sociodemographic variables. We find no heterogeneity in responses with respect to the number of acts of violence committed in the metro. In addition, heterogeneity tests with respect to the intensity of exposure to the scandal (that is, media coverage) yield that the decrease in stated perceived safety grows with the intensity of the press coverage in the country. Both results indicate that the response channel through which the change in stated responses was triggered is purely an informational one, with respondents shifting their stated perception of safety under the influence of information irrespective of the state of the metro network.

This study analysed unique datasets on customer satisfaction with buses and metros as well as on metro performance, which contributes to the existing literature by showing statistical evidence of links between the level of information available and the evaluation of safety by gender. Results support the conclusion that revealed preferences are only revealed up to a certain extent. A change in the context (*e.g.*, an informational shock), can alter perceptions and in turn, can encourage individuals to disclose lower (or higher) satisfaction levels as perceptions of the norm change.

Our analysis point towards the conclusion that stated preferences are subject to variations with respect to informational shocks. More precisely, the mean of revealed preferences highly depends on the context to which a group allows themselves to disclose their tastes and preferences. These results suggest grounds for intervention for service providers who can use informational campaigns to increase customer satisfaction.

7 Acknowledgments

References

Ait Bihi Ouali, L. (2019). Top income tax evasion and redistribution preferences: Evidence from the Panama Papers. Technical report, Aix-Marseille School of Economics, Marseille, France.

Ait Bihi Ouali, L., D. J. Graham, M. Trompet, and A. Barron (2019). Gender differences in the perception of safety in public transport. *Available at SSRN*.

Alesina, A. and P. Giuliano (2011). Preferences for redistribution. In J. Benhabib, A. Bisin, and M. O. Jackson (Eds.), *Handbook of social economics*, Volume 1A, pp. 93–132. Amsterdam: North-Holland.

Anderson, N. H. (1981). *Foundations of information integration theory*, Volume 578. Academic Press New York.

Bénabou, R. and J. Tirole (2002). Self-confidence and personal motivation. *The Quarterly Journal of Economics 117*(3), 871–915.

Bénabou, R. and J. Tirole (2011). Identity, morals, and taboos: Beliefs as assets. *The Quarterly Journal of Economics 126*(2), 805–855.

Cruces, G., R. Perez-Truglia, and M. Tetaz (2013). Biased perceptions of income distribution and preferences for redistribution: Evidence from a survey experiment. *Journal of Public Economics 98*, 100–112.

De Neve, J.-E., C. Imbert, M. Luts, J. Spinnewijn, and T. Tsankova (2017). Behavioural insights and tax compliance: Evidence from large-scale field trials in belgium.

Dolan, R. J. and T. Sharot (2011). *Neuroscience of preference and choice: cognitive and neural mechanisms*. Academic Press.

EN13816 (2002). Transportation − Logistics and Services − Public Passenger Transport − Service Quality Definition, Targeting and Measurement (04), European Parliament and the Council, 2007. Regulation (EC) N. 1371/2007 of the European Parliament and of the Council of 23 October 2007 on rail passengers' rights and obligations. *Off. J. Eur. Union 50* (L 315), 14–41.

Evans, J. S. B. and D. E. Over (2013). *Rationality and reasoning*. Psychology Press.

Festinger, L. (1957). An introduction to the theory of dissonance. *A theory of cognitive dissonance*, 1–31.

Gu, P., M. Triana, T. Rabl, and S.-Y. Byun (2019). # metoo? the role of organizational ethics in employee silence reactions to sex discrimination. In *Academy of Management Proceedings*, Volume 2019, pp. 13030. Academy of Management Briarcliff Manor, NY 10510.

Jorgensen, L. J., G. D. Ellis, and E. Ruddell (2013). Fear perceptions in public parks: Interactions of environmental concealment, the presence of people recreating, and gender. *Environment and Behavior 45*(7), 803–820.

Kahan, D. M., D. A. Hoffman, D. Braman, D. Evans, and J. J. Rachlinski (2012). "they saw a protest": Cognitive illiberalism and the speechconduct distinction. *Stanford Law Review*, 851–906.

Kunda, Z. (1987). Motivated inference: Self-serving generation and evaluation of causal theories. *Journal of personality and social psychology 53*(4), 636.

Kunda, Z. (1990). The case for motivated reasoning. *Psychological bulletin 108*(3), 480.

Kuziemko, I., M. I. Norton, E. Saez, and S. Stantcheva (2015). How elastic are preferences for redistribution? evidence from randomized survey experiments. *American Economic Review 105*(4), 1478–1508.

Margalit, Y. (2013). Explaining social policy preferences: Evidence from the great recession. *American Political Science Review 107*(1), 80–103.

Nyhan, B. and J. Reifler (2010). When corrections fail: The persistence of political misperceptions. *Political Behavior 32*(2), 303–330.

Nyhan, B., J. Reifler, and P. A. Ubel (2013). The hazards of correcting myths about health care reform. *Medical care*, 127–132.

Orgad, S. and R. Gill (2019). Safety valves for mediated female rage in the # metoo era. *Feminist Media Studies 19*(4), 596–603.

Plan International (2018). Girls' safety across the world – summary. *Expert Survey*.

Rabin, M. and J. L. Schrag (1999). First impressions matter: A model of confirmatory bias. *The quarterly journal of economics 114*(1), 37–82.

Sheikh, J. A. and Z. Fayyaz (2019). #metoo: An app to enhancing women safety. In T. Z. Ahram and C. Falcão (Eds.), *Advances in Usability, User Experience and Assistive Technology*, Cham, pp. 546–553. Springer International Publishing.

Sides, J. (2011). Stories, science, and public opinion about the estate tax. *George Washington University Department of Political Science*.

Steinmetz, N. M. and D. M. Austin (2014). Fear of criminal victimization on a college campus: A visual and survey analysis of location and demographic factors. *American Journal of Criminal Justice 39*(3), 511–537.

Taber, C. S. and M. Lodge (2006). Motivated skepticism in the evaluation of political beliefs. *American Journal of Political Science 50*(3), 755–769.

Tetlock, P. E. (1983). Accountability and the perseverance of first impressions. *Social Psychology Quarterly*, 285–292.

A   Appendix - Descriptive Statistics

	Transport Mode: Metro					
	Years					
	2014	2015	2016	2017	2018	2014/18
Women	0.46	0.43	0.47	0.49	0.46	0.46
	(0.50)	(0.50)	(0.50)	(0.50)	(0.50)	(0.50)
Age						
Less than 18	0.04	0.04	0.04	0.04	0.04	0.04
	(0.19)	(0.19)	(0.20)	(0.19)	(0.20)	(0.19)
18-29	0.42	0.39	0.40	0.41	0.37	0.40
	(0.49)	(0.49)	(0.49)	(0.49)	(0.48)	(0.49)
30-39	0.24	0.26	0.24	0.23	0.23	0.24
	(0.43)	(0.44)	(0.43)	(0.42)	(0.42)	(0.43)
40-49	0.14	0.15	0.15	0.15	0.16	0.15
	(0.35)	(0.36)	(0.35)	(0.36)	(0.36)	(0.36)
50-65	0.13	0.14	0.14	0.15	0.15	0.14
	(0.34)	(0.34)	(0.35)	(0.35)	(0.36)	(0.35)
More than 65	0.03	0.03	0.03	0.03	0.04	0.03
	(0.17)	(0.16)	(0.17)	(0.17)	(0.20)	(0.18)
Frequency of use						
Very Often	0.62	0.64	0.63	0.64	0.61	0.63
	(0.49)	(0.48)	(0.48)	(0.48)	(0.49)	(0.48)
Often (min:3times/week)	0.20	0.19	0.19	0.19	0.20	0.19
	(0.40)	(0.39)	(0.39)	(0.39)	(0.40)	(0.39)
Sometimes (min:once/week)	0.10	0.09	0.10	0.09	0.11	0.10
	(0.30)	(0.29)	(0.30)	(0.29)	(0.31)	(0.30)
Rarely (min:once/month)	0.05	0.05	0.05	0.05	0.05	0.05
	(0.22)	(0.22)	(0.22)	(0.22)	(0.22)	(0.22)
Very rarely (less than once/month)	0.03	0.03	0.03	0.03	0.03	0.03
	(0.16)	(0.16)	(0.17)	(0.17)	(0.17)	(0.17)
Most frequent travel motive						
Work	0.59	0.61	0.60	0.57	0.56	0.58
	(0.49)	(0.49)	(0.49)	(0.50)	(0.50)	(0.49)
Education	0.07	0.06	0.07	0.08	0.06	0.07
	(0.25)	(0.24)	(0.25)	(0.27)	(0.24)	(0.25)
Shopping	0.14	0.13	0.13	0.17	0.14	0.14
	(0.34)	(0.34)	(0.33)	(0.37)	(0.35)	(0.35)
Leisure	0.14	0.13	0.14	0.12	0.14	0.13
	(0.35)	(0.33)	(0.34)	(0.32)	(0.35)	(0.34)
Doctor	0.05	0.06	0.05	0.05	0.07	0.06
	(0.22)	(0.24)	(0.22)	(0.22)	(0.25)	(0.23)
Other	0.01	0.01	0.02	0.02	0.02	0.02
	(0.12)	(0.10)	(0.13)	(0.12)	(0.14)	(0.12)

Standard errors in parentheses.
Source: Comet/Nova data.

Table A.1: Descriptive Statistics - Metro Data

	Post MeToo Scandal	
	Post MeToo = 0	Post MeToo = 1
Female = 0	53%	54%
Female = 1	47%	46%

Table A.2: Descriptive Statistics: Distribution of observations for treated and untreated across the cutoff date

B Appendix - Complementary Estimations

	Trains secure	Stations secure	Overall Satisf.	Likelihood Reco.
	(1)	(2)	(3)	(4)
Female (Yes=1)	-0.393***	-0.363***	-0.0418**	-0.00276
	(0.0185)	(0.0191)	(0.0210)	(0.0208)
Post Me Too (Yes=1)	0.119**	0.197***	0.417***	0.284***
	(0.0476)	(0.0473)	(0.0701)	(0.0499)
Post Me Too & Female (Yes=1)	**-0.0752****	**-0.119****	**-0.150****	**-0.0702***
	(0.0374)	**(0.0396)**	**(0.0481)**	**(0.0365)**
Observations	93,276	93,240	93,463	92,193
City Controls	Yes	Yes	Yes	Yes
Sociodemographic	Yes	Yes	Yes	Yes
Chi-square p-value	0.00	0.00	0.00	0.00
Adj. Mc Fadden	0.029	0.03	0.081	0.045

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table B.3: Estimates using ordered logit specifications

C Appendix - Heterogeneity Tests

VARIABLES	Stations are secure for me (1)	Trains are secure for me (2)	Overall Satisfaction (3)
Reference category: Often			
Post & Female & Very Often (Yes=1)	0.00307	-0.0346	-0.0836
	(0.0461)	(0.0363)	(0.0660)
Post & Female & Sometimes (Yes=1)	0.0792	0.0813	0.0530
	(0.0775)	(0.0707)	(0.0799)
Post & Female & Rarely (Yes=1)	0.268**	0.251**	0.106
	(0.105)	(0.0996)	(0.111)
Post & Female & Very Rarely (Yes=1)	0.0221	-0.0800	0.0714
	(0.119)	(0.134)	(0.143)
Observations	93,276	93,240	93,463

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table C.4: Travel frequency heterogeneity effects across sociodemographic characteristics on secure stations – Source: CSS data (Comet/Nova)

VARIABLES	Stations are secure for me (1)	Trains are secure for me (2)	Overall Satisfaction (3)
Reference category: Work			
Post & Female & Education (Yes=1)	0.112	0.204**	0.175
	(0.101)	(0.0985)	(0.116)
Post & Female & Shopping (Yes=1)	-0.00438	-0.0328	-0.0510
	(0.0516)	(0.0596)	(0.0720)
Post & Female & Leisure (Yes=1)	0.0469	0.0920	0.112
	(0.0554)	(0.0620)	(0.0691)
Post & Female & Doctor (Yes=1)	0.163**	0.239***	0.120
	(0.0708)	(0.0678)	(0.0897)
Post & Female & Other (Yes=1)	-0.198	-0.239**	-0.266**
	(0.125)	(0.110)	(0.126)
Observations	93,276	93,240	93,463

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table C.5: Main Travel Purpose heterogeneity effects across sociodemographic characteristics on secure stations – Source: CSS data (Comet/Nova)