Gender Differences in the Perception of Safety in Public Transport

Laila Ait Bihi Ouali[1], Daniel J. Graham[2], Alexander Barron[2], and Mark Trompet[2]

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

Accepted Version [04-February-2020]

Abstract

Concerns over women's safety on public transport systems are commonly reported in the media. In this paper we develop statistical models to test for gender differences in the perception of safety and satisfaction on urban metros and buses using large-scale unique customer satisfaction data for 28 world cities over the period 2009 to 2018. Results indicate a significant gender gap in the perception of safety, with women being 10% more likely than men to feel unsafe in metros (6% for buses). This gender gap is larger for safety than for overall satisfaction (3% in metros and 2.5% in buses), which is consistent with safety being one dimension of overall satisfaction. Results are stable across specifications and robust to inclusion of city-level and time controls. We find heterogeneous responses by sociodemographic characteristics. Data indicates 45% of women feel secure in trains and metro stations (respectively 55% in buses). Thus the gender gap encompasses more differences in transport perception between men and women rather than an intrinsic network fear. Additional models test for the influence of metro characteristics on perceived safety levels and find that that more acts of violence, larger carriages, and emptier vehicles decrease women's feeling of safety.

Keywords: Gender; Safety; Public Transport; Metros; Buses; Customer Satisfaction; Behavioural Responses

1 Introduction

Customer satisfaction and passenger safety and security are priorities for most public transport providers. Aside from the obvious moral, social and equity concerns; lower safety levels are associated with reductions in passenger ridership and thus revenues (*e.g.* Lynch and Atkins 1988, Carter 2005, Delbosc and Currie 2012). This provides ample incentive for operators to intervene by manipulating the public transport characteristics to raise customer satisfaction (Zelinka and Brennan 2001). But safety measures can have different impacts on women relative to men (Yavuz and Welch 2010), and existing research finds that there can be significant mismatches between female customers' safety needs and strategies implemented by public transport operators (Loukaitou-Sideris and Fink 2009). The introduction of women-only carriages is a case in point which has been widely trialled, but has received mixed reactions from female passengers.

A well established strand of literature finds that perceived safety in urban areas varies by gender. Women typically have lower perceived safety in public spaces and a greater fear of being alone in concealed spaces and with strangers (O'Brien 2005, Jorgensen et al. 2013, Steinmetz and Austin 2014, Madge 1997). This points to the conclusion that women may be on average more likely to perceive public spaces as less safe (Pain 1997). However, men are also found to have similar fears, but perhaps under-report in surveys partly to abide by social norms (Yavuz and Welch 2010). For these reasons, it is useful to explore gender differences in the perception of safety rather than women's absolute perception of safety.

While womens' fear of crime in public spaces has been widely studied (*e.g.* Hall 1985, Riger and Gordon 1989, Valentine 1990, Gilchrist et al. 1998, Koskela and Pain 2000, Pain 2001), safety in public transport has received little statistical attention. Does the perception of safety in public transport differ by gender, and to what extent? If so, what are the potential drivers of these differences? These are the questions we seek to answer in this paper.

This paper focuses on the individuals' perception of their own security. We define it as the estimated risk or threat of an intentional personal attack or aggression[1]. Dependent variables used in this study are ordinal thus our empirical strategy relies on ordered probit specifications. To the best of our knowledge, this study is the first cross-country statistical analysis of gender differences in the perception of both safety and satisfaction in public transport. The first analysis uses data from two original Customer Satisfaction Surveys (CSS) respectively conducted for (i) urban metros (over 2014-2018) and (ii) buses (over 2009-2018). The richness and abundance of the data is ideal for our study as it allows us to provide robust results, generate conclusions with wide geographical relevance, and compare gender gaps across different transport modes.

Following quantification of gender differences in the perception of safety in public transport, the second aim of this paper is to disentangle potential drivers of observed differences. This is relevant as it assesses whether the design and other attributes of public transport help improve perceived levels of safety. The motivation behind this analysis is developed in the urban social geography and planning literatures which explain that initiatives to "design out" fear of crime in urban areas have little effects on perceived safety (Koskela and Pain 2000). We draw on Key Performance Indicators data (KPI) that covers extensively 25 metro systems for 2014-2018.

Results suggest a significant gap between men and women both in perceived satisfaction and safety. We find that women are 10% more likely than men to feel unsafe in metros and 6% more likely to feel unsafe in buses. This gender gap is smaller for overall satisfaction, as women are 3% more likely to be dissatisfied with the general service in metros and 2.5% in buses. This smaller magnitude is consistent with safety being only one important dimension of overall satisfaction (Oliver 1997). Despite this gap, women are on average satisfied with both safety and the service. Results are stable across specifications and robust to city and

[1]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* and *security* identically.

time controls. Our results show heterogeneity with age and by geographical area. On the influence of metro characteristics on safety levels, we find that more acts of violence, larger carriages, and emptier vehicles decrease the feeling of safety among women.

2 Data

The data used in this study have been collected annually over the period 2009 to 2018 via responses to Customer Satisfaction Surveys (CSSs) from users of urban metros and buses. 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, and with bus operators through the IBBG (International Bus Benchmarking Group). The CoMET and Nova groups cover 25 cities across Europe, Americas and Asia, while the IBBG group is a consortium of 14 bus networks across Europe, Asia and North America. The detailed composition of each consortium is presented in Appendix A.

Both the metro and bus CSSs have a similar structure (Trompet et al. 2013, 2018). The first part of the questionnaire contains statements relating to eight customer service areas defined under European Norm 13816: availability, time, information, comfort, security, 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. 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-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 an increasing scale going from "1−Agree Strongly" to "5–Disagree Strongly". Questions on safety in transport follow this structure. The scale of possible responses on the perception of overall satisfaction is similar and ranges increasingly from "1−Strongly Dissatisfied" to "5–Strongly Satisfied". The second section of the survey asks respondents to select, in order of preference, the three most important customer service areas 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 and buses. Metro data records a total of 137,513 observations for 25 cities over the period going from 2014 to 2018. Bus data contains a total of 189,890 observations for 14 cities between 2009 and 2018. Descriptive statistics are in Tables 9 and 10 in Appendix respectively for metros and buses. The distribution of genders is rather equal for both datasets for all years; in addition, respondents are often daily commuters who travel 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.

The CSS dataset's main advantage is in the abundance of observations: data covers a critical number of cities worldwide over a consequent time period, which conveys our results robustness and external validity of our results. Another unique advantage is in the representativeness of the data as transport providers sample their respondents every year that way. This cross-sectional dataset does not follow respondents over time – yet, the representativeness of the data over each year allow us to assume that characteristics do not change much over time.

Smartcard datasets often contain more information on individuals, such as the metro line they take, and OD pairs can enable us to pinpoint daily commuters and infer where they live. However, despite the lack of side variables of this sort, survey data still remains until today the only resource that can tell us about perceptions and customer satisfaction. Therefore, survey data still remains essential to complement demand analysis and observe behaviours through the prism of satisfaction surveys.

3 The Model

All outcome 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 the 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 scale is the same for the outcome variable measuring overall satisfaction with the service, but ranges from “Strongly Dissatisfied” to “Strongly Satisfied”.

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 $\Pr[y_i \leq j|x] = \phi(\alpha_j - x_i'\beta)$ 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 dependent 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

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

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

$$\begin{aligned} Pr(y_i = j) &= Pr(\alpha_{j-1} < y_i^* \leq \alpha_j) \\ &= Pr(\alpha_{j-1} < x_i\beta + u_i \leq \alpha_j) \\ &= Pr(\alpha_{j-1} - x_i\beta < u_i \leq \alpha_j - x_i\beta) \\ &= \phi_{ij} - \phi_{ij-1} \end{aligned} \tag{2}$$

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.

Covariates representing sociodemographic characteristics in Equation 1 contribute to a better understanding of the drivers of perceived safety. Our covariates also helps distinguishing groups of people who are more or less likely to perceive a high (or low) level of safety. First, we include measures of travel frequency and travel purpose (*e.g.* work, education). These two dimensions characterise several profiles and whether travel is constrained or for travel at different dates, times and frequencies. We also add time dummies to account for potential exogenous shocks and evolutions over time. Finally, city fixed effects are deemed helpful to account for network (buses or metro) characteristics, city structure and safety levels and also country civic values[2].

The choice of an ordered probit over an ordered logit model relies on the assumption that the underlying distribution of the observed outcome is normal. In an ordered probit specification, the observed outcome (e.g., "Agree") reflects a threshold is met for the underlying latent variable which is normally distributed. The plotted distribution of satisfaction in our data provided a relatively normal distribution with a mean satisfaction at about scale 4 (and of 3.25 for both security in trains and security in stations). Thus, normal distributions appear more suited to define satisfaction levels. Finally, when it comes to the computation of marginal effects, both logit and probit models make essentially similar predictions[3].

Ordered probit models rely on a couple of assumptions that appear as quite strong: (i) the constant threshold assumption; (ii) the distributional assumption. The distributional assumption assumes there is no additional individual heterogeneity between individual realisations. In our case, women may have heterogeneous behaviours as unmeasured variables can affect the chances of feeling more or less safe in public transport (*e.g.*, risk aversion). Therefore, to counteract this limitation, we compute multilevel mixed-effects ordered probit models that account for this individual heterogeneity (see Tables 16 and 11 for random intercept and random coefficient specifications).

Mixed-effects ordered probit models contain both fixed and random effects. As such, they

[2]Small differences between bus and metro surveys exist as each group sets their own rules within our framework. This is at the root of the differences in the "Main Travel Purpose" variable: the bus consortium grouped 'Education' with 'Work' as they considered them as similar types of public transport uses.

[3]We computed estimates using logit models. Results are widely comparable in sign and significance.

allow for many levels of nested clusters of random effects. Using the original framework presented in equation (2), we now assume a series of K independent clusters. This method accounts for unobserved heterogeneity at the city level. Therefore, conditional on a set of fixed effects x_{ik}, and a set of random effects u_k, we can derive the probability of observing outcome j as

$$\begin{aligned} Pr(y_{ik} = j|\alpha, u_k) &= Pr(\alpha_{j-1} < y_i^* k \leq \alpha_j) \\ &= Pr(\alpha_{j-1} < x_{ik}\beta + z_{ik}u_k + e_{ik} \leq \alpha_j) \end{aligned} \tag{3}$$

where y_{ik}^* is the latent continuous response (*i.e.*, a statement on safety or satisfaction in public transport). e_{ik} is the error term which follows a standard normal distribution and is independent of u_k.

This multilevel mixed effect model accounts for city-level heterogeneity which is due to the nature of our dataset. The Customer Satisfaction Survey (CSS) is a cross-sectional dataset at the individual level. As such, one observation line represents the level of satisfaction at time t for a given individual i in a given city. Therefore, the city level is the smallest level at which we have unobserved heterogeneity we can control for. We acknowledge that only city level heterogeneity is accounted for, not individual level heterogeneity. However, since random effects cannot adjust for confounding in any case, this should not materially affect the overall conclusions drawn from this study.

Coefficients remain of the same order of magnitude and indicates that in spite of its strength, this assumption does not impair the quality of our estimations. Also, although responses can differ by subgroup, the aim of this paper is essentially to observe the difference in perception on a worldwide scale while clearing effects of socio-demographics (age, habits) and allows us to treat satisfaction levels as a non-linear scale, which is a crucial condition.

A second limitation of ordered probit models is contained in marginal probability effects whose signs' change once when moving from the smallest to the largest outcome. We accept this assumption as it is compatible with the fact that the satisfaction scale, despite being non-linear, is an increasing one.

4 Results

4.1 Base Results

4.1.1 Quantification of the gender effect

Table 1 below presents ordered probit estimates for safety-related questions, estimated with covariates for sociodemographic characteristics included along with time and city fixed effects. Dependent variables are the discrete representation of statements on safety and overall satisfaction. Sensitivity tests show that all estimates, both for buses and metros, are stable across specifications and robust to covariate inclusion.

Table 1 shows a negative and significant gap by gender for both metros and buses. This gap is observed for both safety and overall satisfaction statements. This corroborates the urban economics and crime literature (Hall 2005, Riger and Gordon 1989) as women's bigger fear

	Metros			Buses	
	The train is a secure place for me	Stations are a secure place for me	How satisfied are you overall with the *metro* service?	The bus is a secure place for me	How satisfied are you overall with the *bus* service?
Female (Yes=1)	-0.235***	-0.235***	-0.0452***	-0.105***	-0.0372***
	(0.00539)	(0.00539)	(0.00548)	(0.00516)	(0.00515)
P-value of coeft. female	<0.001	<0.001	<0.001	<0.001	<0.001
McFadden Adj. R^2	0.043	0.045	0.081	0.019	0.029
Observations	169,582	169,658	169,831	179,773	180,208

Standard errors in parentheses.
*** p<0.01, ** p<0.05, * p<0.1
Source: Customer Satisfaction Surveys, Comet/Nova and IBBG, own calculations.
Note: All estimates from regressions computed above contain city and year dummies.

Table 1: Ordered Probit Estimates: Gender Coefficients on Perception of Safety

of crime in public spaces is also now observed for public transport. This result also extends this literature (Twinam 2017, Phillips and Sandler 2015) by indicating that women and men differ not only in propensity to be targeted by crime but also in their perceptions. In addition, since the data covers a wide range of cities worldwide and results are robust, this observation is applicable to cities with a strong external validity.

These first results highlight a first contribution of our analysis. The issue of women's safety in public spaces has been included in the public discourse since the 1980s (Shaw 2002). However, the literature has been limited thus far by a lack of no homogeneous, standardised way to record violences against women, as data is often incomplete or inaccurate and is subject to under-reporting (UN Women 2016). Yet we now find, using a homogenous dataset recording perceptions at a worldwide scale, that perceptions of safety are commonly share across the globe and that the gap in perception is significant. However, the magnitude of coefficients presented in Table 1 cannot be interpreted directly - marginal effects are presented and extensively discussed in the rest of this section. Yet, these coefficients are still relevant. They are negative and significant at the 1% level: as such, they show there is a significant gap in the perception of safety between men and women.

Assessing the magnitude of the difference between men and women in the perception of safety requires the computation of marginal effects. The interpretation of the marginal effects is straightforward for continuous variables, with a one unit change in the explanatory variable resulting in an increase or decrease in the probability equal to the size of the marginal effect (all other things equal). Since y^* cannot be observed and is purely artificial, its interpretation is not of interest.

The most natural way to interpret ordered response models (and discrete probability models in general) is to determine how a marginal change in one regressor changes the distribution of the outcome variable, *i.e.* all the outcome probabilities. The main focus in the analysis of ordered data should be put on the conditional cell probabilities given by

$$Pr[y = j|x] = F(\mu_j - x'\beta) - F(\mu_{j-1} - x'\beta). \tag{4}$$

with F being the variance of the distribution function and β the vector of coefficient attached to the vector of observable variables x. In order to identify the parameters of the model we have to fix location and scale of the argument in F, the former by assuming that x does not contain a constant term, the latter by normalizing the variance of the distribution function F

$$MPE_{jl}(x) = \frac{\delta Pr[y = j|x]}{\delta x_l}. \tag{5}$$

The marginal effects represent the variation in the probability of picking one given response (*e.g.* “Agree”) if the individual is a woman. Marginal effects are computed for every possible response ranging from “Strongly Disagree” to “Strongly Agree”. Figure 1 indicates that women are 4% more likely than men to “Disagree Strongly” with the statement that they feel secure in trains. Marginal effects for gender are presented in Figures 1 and 2 below.

In the case of a binary variable, the marginal effect is the change in predicted probability based on whether a respondent falls into that category or not. When calculating marginal effects all remaining variables assume their respective average values. As such, the marginal effects show the change in the predicted probability for each gender for an average respondent, according to the variable being considered.

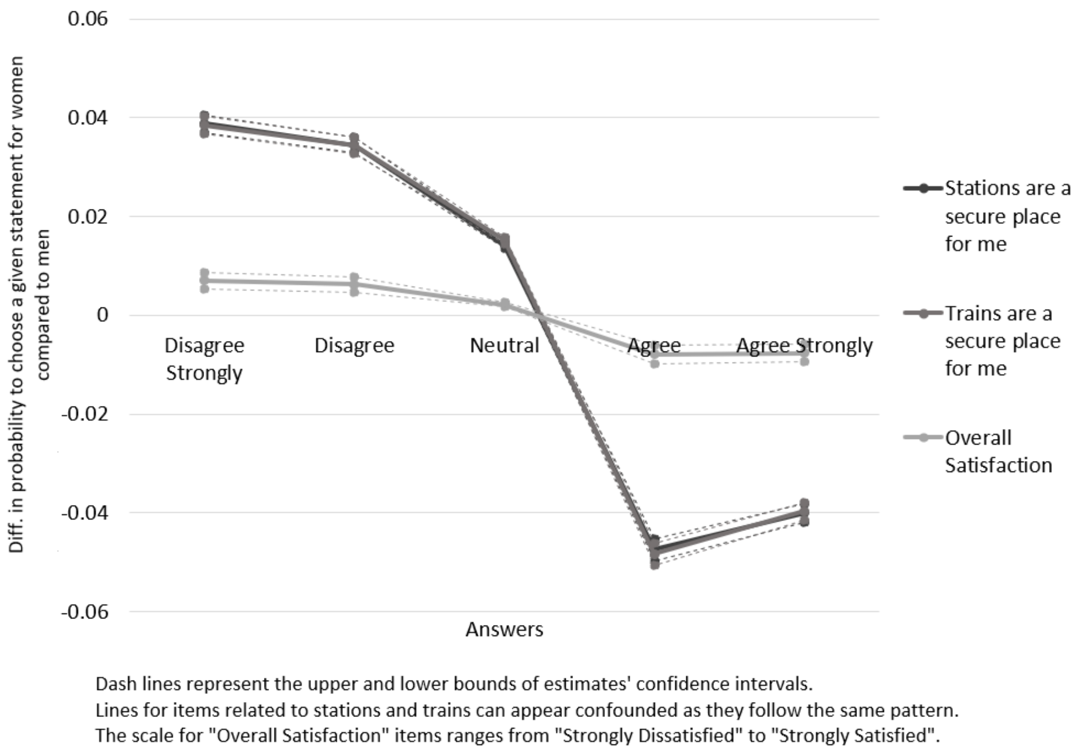
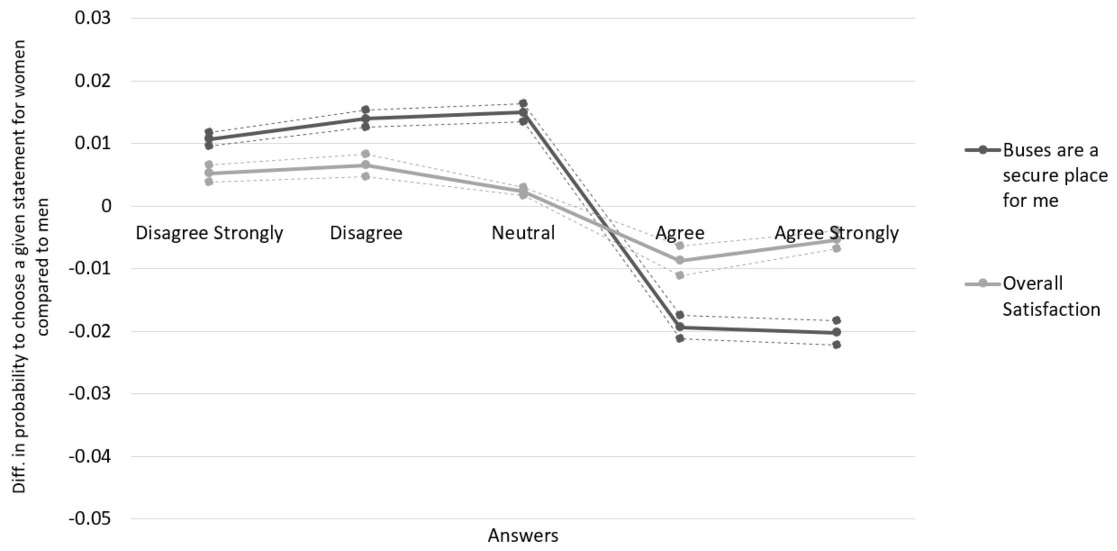


Figure 1: Marginal Effects for Safety-Related Questions - Metro Data



Dash lines represent the upper and lower bounds of estimates' confidence intervals.
The scale for "Overall Satisfaction" items ranges from "Strongly Dissatisfied" to "Strongly Satisfied".

Figure 2: Marginal Effects for Safety-Related Questions - Bus Data

Figures 1 and 2 shows that overall, women are 10% more likely to make a negative statement on security than men. This figure shows that the two statements on security present marginal effects of a similar magnitude, which is in line with the fact that these statements are very close both in label and in nature. However, we note that we find a much larger effect magnitude for statements regarding safety compared to overall satisfaction. For comparison, Tables 12 and 13 in the Appendix shows gender differences for all other dimensions of customer satisfaction available in the Customer Satisfaction Survey (*e.g.* Availability, Information, Customer Care), respectively for metros and buses. Results indicate that there is a negative and significant gender gap in the perception of all dimensions of customer satisfaction. However, the magnitude of the gender gap is much larger for safety than it is for any other dimensions of customer satisfaction. These results are indicate that safety is the main vector of gender differences in perception.

While we observe this gender gap, the data also reveal that women are generally happy with the overall quality of service. Figures 8 and 9 in the Appendix show the general probability of women in picking a given answer for each statement. These results suggest that although the gap between men and women is significant, the levels of satisfaction and safety of female customers still remains reasonably high.

4.1.2 Analysis of other socio-demographic variables

Table 2 below shows coefficients for the socio-demographic and economic variables for both transport modes. Stated feelings of safety in metros appear to decrease with age, although this feeling of safety is mostly increasing with age for buses. However, satisfaction with the general service evolves similarly for buses and metros and appears to increase from 40 years old onwards. Responses based on the main travel purpose appear to evolve similarly for both buses and metros: people travelling mostly for shopping and leisure purposes declare significantly higher safety levels than people who travel mostly to go to work or school. Individuals who travel for leisure or shopping are more likely to travel on their own terms, pick their destinations and routes as well as travel dates and times; they are also more likely to travel with people they know. In contrast, travelling for work or to go to school is more likely to abide by constraints.

Travel frequency is also highly correlated with stated levels of safety, although perceptions by subgroups are different for buses than for metros. In both cases, individuals who travel "often" feel safer than those who travel "very often". On the other hand, individuals using public transport "Very Rarely" are more likely to feel less safe.

	Metro			**Buses**	
	The train is a secure place for me	**Stations are a secure place for me**	**How satisfied are you overall with the service?**	**The bus is a secure place for me**	**How satisfied are you overall with the service?**
Female (Yes=1)	-0.235***	-0.235***	-0.0452***	-0.105***	-0.0372***
	(0.00539)	(0.00539)	(0.00548)	(0.00515)	(0.00515)
Age (ref: 18-29)					
Less than 18	0.238***	0.197***	0.359***	0.241***	0.426***
	(0.0144)	(0.0144)	(0.0148)	(0.0126)	(0.0127)
30-39	-0.0279***	-0.00773	-0.0442***	0.0329***	-0.0666***
	(0.00703)	(0.00703)	(0.00716)	(0.00692)	(0.00690)
40-49	-0.00685	0.0141*	0.0556***	0.0701***	0.0217***
	(0.00826)	(0.00826)	(0.00841)	(0.00784)	(0.00783)
50-65	0.0251***	0.0505***	0.152***	0.0857***	0.0984***
	(0.00857)	(0.00857)	(0.00875)	(0.00771)	(0.00771)
Over 65	0.152***	0.178***	0.325***	0.212***	0.282***
	(0.0158)	(0.0158)	(0.0162)	(0.0133)	(0.0133)
Main travel purpose (ref: Work)					
Education	0.0497***	0.0451***	0.0761***		
	(0.0116)	(0.0116)	(0.0118)		
Shopping	0.0963***	0.0866***	0.155***	0.0421***	0.197***
	(0.00853)	(0.00853)	(0.00869)	(0.0134)	(0.0134)
Leisure	0.159***	0.151***	0.252***	0.104***	0.231***
	(0.00994)	(0.00994)	(0.0102)	(0.00833)	(0.00835)
Doctor	0.0614***	0.0378***	0.189***	-0.135***	0.0448***
	(0.0128)	(0.0128)	(0.0131)	(0.0170)	(0.0170)
Other	0.0445**	0.0296	0.128***	0.0143	0.0237
	(0.0211)	(0.0211)	(0.0215)	(0.0158)	(0.0157)
Frequency use trains (ref: very often)					
Often	0.0836***	0.0696***	0.128***	0.0717***	0.115***
	(0.00733)	(0.00734)	(0.00748)	(0.00653)	(0.00653)
Sometimes	0.0640***	0.0394***	0.155***	0.0811***	0.116***
	(0.0104)	(0.0104)	(0.0106)	(0.00880)	(0.00881)
Rarely	0.0541***	0.0231*	0.198***	0.00361	0.0131
	(0.0134)	(0.0134)	(0.0138)	(0.0129)	(0.0129)
Very Rarely	0.00238	-0.0320*	0.170***	-0.145***	-0.214***
	(0.0171)	(0.0171)	(0.0176)	(0.0169)	(0.0169)
P-value of coeft. female	<0.001	<0.001	<0.001	<0.001	<0.001
Observations	169,582	169,658	169,831	179,773	180,208
McFadden Adj. R^2	0.043	0.045	0.081	0.019	0.029

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1
Source: Customer Satisfaction Surveys, Comet/Nova and IBBG, own calculations.
Note that for buses, the reference category for the main travel purpose is Work/Education
Note: All estimates from regressions computed above contain city and year dummies.

Table 2: Estimates for sociodemographic variables - Bus and Metro Data

4.2 Heterogeneity Checks

4.2.1 Heterogeneous responses by area

Figures 3, 4, and 5 show the marginal coefficients of the gender gap by geographic region. This enables us to assess whether there is heterogeneity at the continent level.

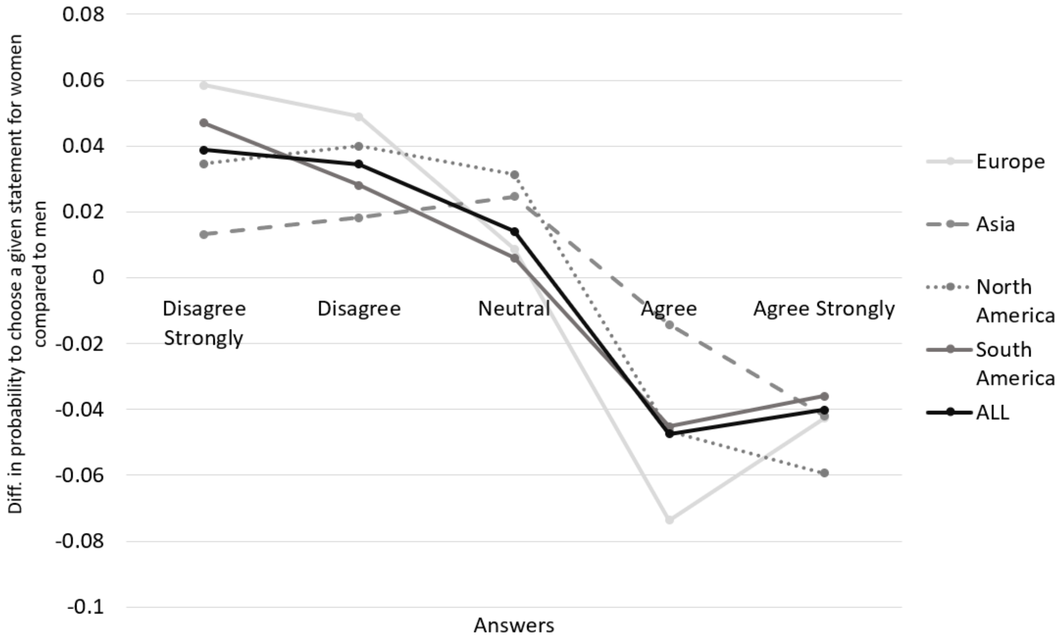


Figure 3: Regional Marginal Effects for Safety-Related Statement: "Stations are a secure place for me"

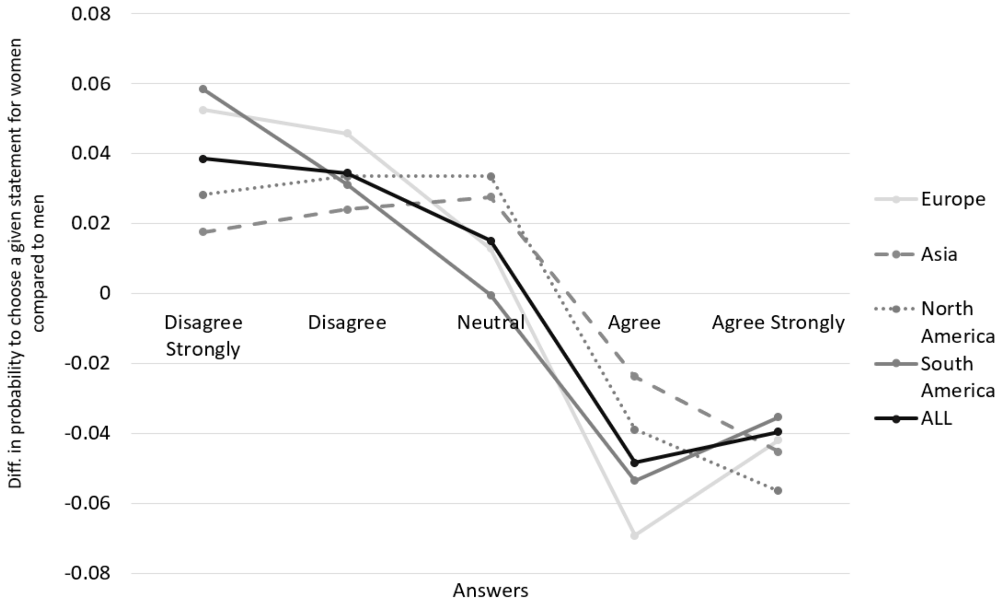


Figure 4: Regional Marginal Effects for Safety-Related Statement: "Trains are a secure place for me"

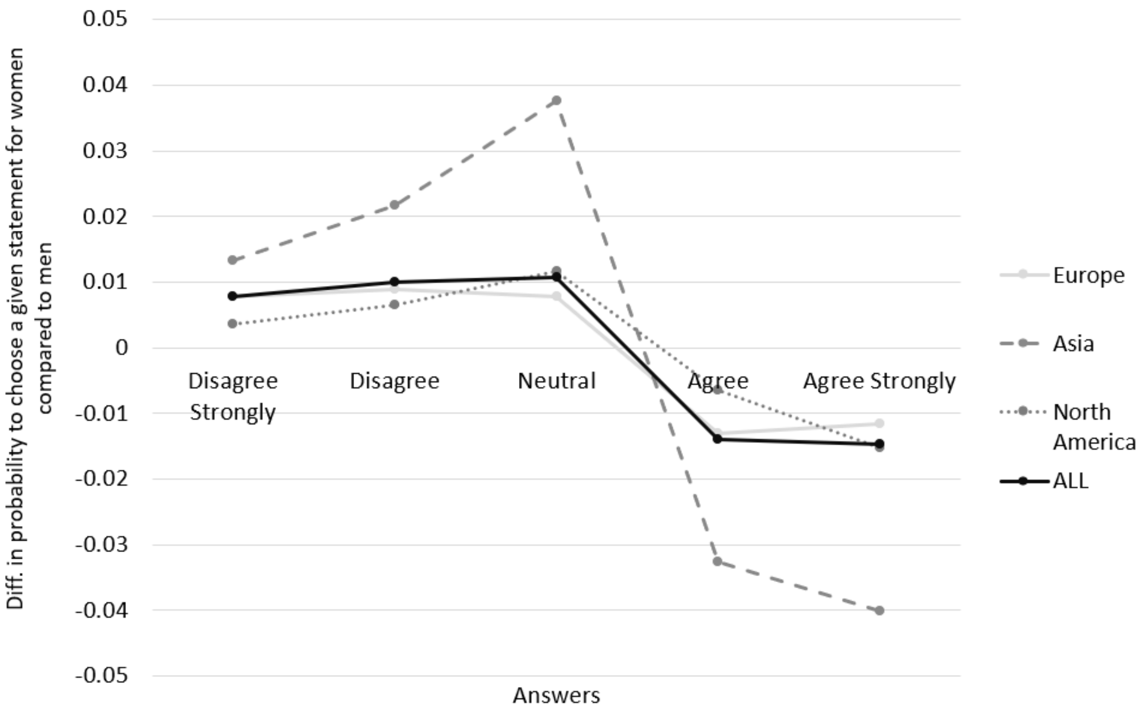


Figure 5: Regional Marginal Effects for Safety-Related Statement: "The bus is a secure place for me"

To preserve anonymity and respect the confidential nature of our data we aggregated the information to the continent level. We find that for safety statements the evolution is the same for all continents but with a different magnitude by continent. Women are found have negative experiences in public transport such as harassment (Condon et al. 2007, European Commission 2014): our results suggest that they also share the same perceptions of safety. For metros, results suggest that both Europe and North America have a bigger gender gap in satisfaction compared to Asia or South America. For buses, we observe that Europe and North America have a comparable gender gap, with Asia presenting more differences and a higher propensity of women to be "Neutral" regarding their perception of safety. This homogeneity is not observable for the "overall satisfaction" statement (see Figures 6 and 7 in Appendix).

4.2.2 Heterogeneous responses by age

Table 3 presents results on the heterogeneity of the gender effect by age. We find that overall satisfaction with public transport increases with age and that satisfaction increases at an increasing rate with age for safety statements. We find that the perception of safety in buses also increases with age. However, this effect is not observed for metros, where the perception of safety decreases at a decreasing rate with age. Since younger women are over represented in our data relative to the population, our estimated effect is likely conservative.

	Metro			Buses	
	The train is a secure place for me	Stations are a secure place for me	How satisfied are you overall with the service?	The bus is a secure place for me	How satisfied are you overall with the service?
Female (Yes=1)	-0.291***	-0.312***	-0.0405***	-0.145***	-0.0123
	(0.00829)	(0.00830)	(0.00844)	(0.00878)	(0.00877)
Age (ref: 18-29)					
Less than 18	0.272***	0.224***	0.411***	0.300***	0.506***
	(0.0167)	(0.0167)	(0.0171)	(0.0164)	(0.0165)
30-39	-0.0603***	-0.0625***	-0.0624***	-0.00595	-0.0832***
	(0.00936)	(0.00937)	(0.00952)	(0.00991)	(0.00987)
40-49	-0.0626***	-0.0586***	0.0481***	0.0403***	0.0388***
	(0.0112)	(0.0112)	(0.0114)	(0.0113)	(0.0113)
50-65	-0.0498***	-0.0335***	0.179***	0.0334***	0.146***
	(0.0118)	(0.0118)	(0.0120)	(0.0113)	(0.0113)
Over 65	0.0895***	0.0994***	0.381***	0.165***	0.343***
	(0.0202)	(0.0202)	(0.0208)	(0.0179)	(0.0180)
Age # Female					
Female & (age: <18)	-0.192***	-0.178***	-0.211***	-0.169***	-0.197***
	(0.0324)	(0.0324)	(0.0332)	(0.0257)	(0.0259)
Female & (age: 30-39)	0.0685***	0.117***	0.0410***	0.0749***	0.0382***
	(0.0133)	(0.0133)	(0.0135)	(0.0137)	(0.0137)
Female & (age: 40-49)	0.116***	0.151***	0.0166	0.0562***	-0.0295*
	(0.0157)	(0.0157)	(0.0160)	(0.0154)	(0.0154)
Female & (age: 50-65)	0.151***	0.171***	-0.0518***	0.0957***	-0.0803***
	(0.0160)	(0.0160)	(0.0164)	(0.0147)	(0.0147)
Female & (age: over 65)	0.147***	0.184***	-0.135***	0.0949***	-0.115***
	(0.0302)	(0.0302)	(0.0311)	(0.0246)	(0.0247)
P-value of coeft. female	<0.001	<0.001	<0.001	<0.001	0.156
Observations	169,582	169,658	169,831	179,616	180,047
McFadden Adj.R^2	0.043	0.045	0.082	0.019	0.029

Standard errors in parentheses.
*** p<0.01, ** p<0.05, * p<0.1 Source: Customer Satisfaction Surveys, Comet/Nova and IBBG, own calculations.
Note: All estimates from regressions computed above contain city and year dummies.

Table 3: Estimates for heterogeneity based on age - Metro and Bus Data

4.2.3 Heterogeneous responses by travel purpose

Table 4 presents estimates testing for heterogeneity by travel purpose. Results are similar for both statements on safety. There is significant heterogeneity between men and women only for shopping and leisure, with the latter being significantly less satisfied.

	Metro			Buses	
	The train is a secure place for me	Stations are a secure place for me	How satisfied are you overall with the service?	The bus is a secure place for me	How satisfied are you with the service?
Female (Yes=1)	-0.212***	-0.206***	-0.0223***	-0.120***	-0.0415***
	(0.00686)	(0.00686)	(0.00697)	(0.00595)	(0.00594)
Main travel purpose (ref: Work)					
Education	0.0624***	0.0571***	0.0973***		
	(0.0145)	(0.0145)	(0.0148)		
Shopping	0.146***	0.148***	0.173***	0.0140	0.181***
	(0.0112)	(0.0112)	(0.0114)	(0.0202)	(0.0203)
Leisure	0.183***	0.181***	0.286***	0.0695***	0.218***
	(0.0121)	(0.0121)	(0.0124)	(0.0108)	(0.0108)
Doctor	0.0587***	0.0426**	0.220***	-0.150***	0.0721***
	(0.0175)	(0.0175)	(0.0179)	(0.0270)	(0.0270)
Other	0.0812***	0.0597**	0.185***	0.00126	0.0253
	(0.0288)	(0.0288)	(0.0295)	(0.0233)	(0.0233)
Main travel purpose # Female					
Female & Education	-0.0273	-0.0245	-0.0499**		
	(0.0218)	(0.0219)	(0.0223)		
Female & Shopping	-0.103***	-0.128***	-0.0353**	0.0486*	0.0287
	(0.0153)	(0.0153)	(0.0156)	(0.0256)	(0.0258)
Female & Leisure	-0.0515***	-0.0625***	-0.0756***	0.0705***	0.0272*
	(0.0158)	(0.0158)	(0.0162)	(0.0140)	(0.0140)
Female & Doctor	0.00372	-0.0108	-0.0618***	0.0265	-0.0414
	(0.0227)	(0.0227)	(0.0232)	(0.0335)	(0.0336)
Female & Other	-0.0756*	-0.0612	-0.120***	0.0244	-0.0106
	(0.0413)	(0.0413)	(0.0422)	(0.0306)	(0.0306)
P-value of coeft. female	<0.001	<0.001	0.001	<0.001	<0.001
Observations	169,582	169,658	169,831	179,615	180,046
McFadden Adj.R^2	0.042	0.045	0.081	0.019	0.029

Standard errors in parentheses.
*** p<0.01, ** p<0.05, * p<0.1 Source: Customer Satisfaction Surveys, Comet/Nova and IBBG, own calculations.
Note: For the bus data, the reference category for the main travel purpose is "Work/Education".
For metro data, the reference category is "Work" since the main travel purpose categories are more detailed.
Note: All estimates from regressions computed above contain city and year dummies.

Table 4: Estimates for heterogeneity based on travel purpose - Metro and Bus Data

5 Supplementary analyses

The customer satisfaction data we use in this paper originate from an annually repeated cross-sectional survey. So far, our results have been estimated conditional on time and city fixed effects. However, as the underlying data comes from a repeated cross-section, it cannot be assumed that the participating respondents remain identical and constant over time. Our ordered probit estimates help assess the impact of explanatory variables on an ordinal variable. However, ordered probit estimates cannot account for unobserved heterogeneity among respondents over time. In our previous regressions, we try to disentangle the extent to which women's perception of safety differs from that of men via inclusion of socio-demographic covariates for gender, age, and travel conditions (purpose and frequency). But while these are helpful, they do not adjust comprehensively for other potential important time-varying or time–invariant effects, for instance marital status or the number of children. Since such characteristics could be relevant for women' perception of safety, their omission from our model negates a causal interpretation of results.

In order to address this issue, we adopt a pseudo-panel methodology which was first developed by Deaton (1985) for individual level data. Similar to a standard panel methodology, the pseudo-panel approach carefully accounts for unobserved time-invariant heterogeneity facilitating an improved understanding of the causal mechanisms by which outcomes are produced. We adopt the pseudo-panel approach as a means of estimating individual preferences in the absence of true individual panel data. It enables us to get adjust for unobserved heterogeneity among cohorts and to corroborate our previous ordered probit estimates with a different methodology.

The pseudo-panel method restructures the data and forms cohorts that are consistent over time. The cross-sectional data used to construct pseudo-panels must include information on at least one observable and time invariant variable by which observations can be grouped into cohorts. Then, cohort means are computed for all included variables and tracked over time to form a matrix of all cohorts as a pseudo-panel. This pseudo-panel of means does not suffer from attrition.

We use the city identifier and the response date available for the metro customer satisfaction survey to identify the cohorts and generate the mean variables. For the bus data, response dates are not available for the whole period and we therefore construct our psuedo-panel only for metros.

Under this approach the dependent variable represents the satisfaction scale as a continuous variable. It is defined as

$$y_{ct} = x_{ct}\beta + \alpha_c + \delta_t + u_{ct} \tag{6}$$

where c represents a city / response date cohort, t represents the date at which individuals fill the survey, x_{ct} is a vector of observable characteristics, α_c are cohort fixed effects, δ_t are time fixed effects, and u_{ct} the error term. Assuming the size of the cohorts is sufficiently large and the composition relatively stable over the years, the daily cohort average of the firm-specific time-invariant effect can be transformed into a city time invariant effect α_c which allows us to control for unobserved heterogeneity between cohorts.

Each observation in the following analysis is thus the mean individual response of a city cohort at time t, hence allowing us to estimate the effect of the average share of women on the level of safety perceived by individuals.

Table 5 shows the results from our pseudo-panel models. We find a significant and negative gap between men and women in our sample, indicating that women are less likely to feel safe in metros compared to men. This is applicable to both the perception of safety in trains and stations. While it is difficult to compare the magnitude of the coefficients in the Fixed Effects (FE) estimation model with those of the ordered probit model, we note that all coefficients linked to gender have the same negative sign and are highly significant in this specification.

The pseudo-panel method we use relies on the assumption that the covariates are endogeneous and does not adjust for potential time-varying unobserved heterogeneity. We test whether results are robust to these assumption via a Generalized method of moments (GMM) methodology using instrumental variables derived for our pseudo-panel. We compute estimators from two-stage least-squares generalizations of simple panel-data estimators for exogenous variables. In this study, we use the lagged differences of variables as instrumental variables for levels:

$$\begin{aligned} y_{ct} &= x_{ct}\beta + \alpha_c + \delta_t + u_{ct} \\ with\; x_{ct} &= \Delta x_{ct-1}\gamma + \Delta x_{ct-2}\phi + v_{ct} \end{aligned} \tag{7}$$

where Δx_{ct-1} and Δx_{ct-2} represent the vectors of instrumental variables, respectively the first and second lags of the growth of variables; v_{ct} is the error term. Table 6 presents the results from the GMM specification. We find that a larger share of women in a given city decreases the level of the results are robust to this specification. Regarding the magnitude of coefficients, we observe that GMM estimates record that a larger share of women decrease even more the perception of safety compared to the regular pseudo-panel specification. This appears to indicate that not controlling for endogeneity biased estimates upwards.

	The train is a secure place for me		Stations are a secure place for me		How satisfied are you overall with the *metro* service?	
Share of female respondents	-0.216***	-0.200***	-0.147***	-0.138***	-0.00795	-0.000838
	(0.0420)	(0.0428)	(0.0419)	(0.0428)	(0.0480)	(0.0483)
Socio-demographic controls	No	Yes	No	Yes	No	Yes
Observations	2,955	2,899	2,955	2,899	2,954	2,898
R-squared	0.009	0.024	0.004	0.016	0.000	0.043
Number of cities	25	25	25	25	25	25

Standard errors in parentheses.
*** p<0.01, ** p<0.05, * p<0.1
Source: Customer Satisfaction Surveys, Comet/Nova, own calculations.

Table 5: Pseudo Panel Estimates: Gender Coefficients on Perception of Safety - Metros

	The train is a secure place for me		Stations are a secure place for me		How satisfied are you overall with the *metro* service?	
Share of female respondents	-3.243*	-2.948**	-3.615*	-3.236*	-4.887*	-4.489*
	(1.697)	(1.492)	(2.150)	(1.875)	(2.728)	(2.421)
Year dummies	No	Yes	No	Yes	No	Yes
Observations	2,189	2,189	2,189	2,189	2,189	2,189
R-squared	0.11	0.13	0.05	0.06	0.14	0.16
Number of cities	25	25	25	25	25	25

Standard errors in parentheses.
*** p<0.01, ** p<0.05, * p<0.1
Source: Customer Satisfaction Surveys, Comet/Nova, own calculations.

Table 6: GMM Pseudo Panel Estimates: Gender Coefficients on Perception of Safety - Metros

Finally, Tables 11 and 16 in Appendix present results of a multilevel mixed effects model with both a random intercept and a random coefficient. This specification accounts for unobserved heterogeneity at the city level and includes a random coefficient for the gender variable at the city level. Estimates indicate that our results are robust to this more general specification for the safety-related outcomes, which confirms that safety in transport is subject to a gap between men and women.

6 Analysing potential drivers of perceived safety in metros

Having found convincing evidence for a gender gap in public transport safety we now move on to the second aim of this paper, which is to disentangle potential drivers of these observed differences. To do this we use extensive data collected by the TSC over the period 2014 to 2018 for metro systems in 25 cities worldwide. An equivalent version of data is available for buses, but it does not cover the same time span, and is not as complete; therefore, our analysis is restricted to urban metros.

The choice of covariates is based on the physical characteristics in the previous literature that affect individuals' perception of risk and fear (Atkins 1990, Valentine 1990). The metro data allow us to test for the influence of the following covariates on the perceptions of safety: (i) number of staff members in metro stations (both own staff and contractors), (ii) number of cars per train, (iii) the average number of stations served by line (on average), (iv) the number of violent acts committed in the metro, (v) the total car capacity (seating and standing), and (vi) the metro ridership.

The number of staff members encompasses here the metro providers' own staff present in metro stations. Customers can refer to station employees in case of emergency, and they are easy to notice as they potentially have a marker (uniform, badge) that can indicate that they work for the metro company. Therefore, we expect a potential positive or neutral effect of staff members. Indeed, more staffing does not necessarily means constant staffing or staff members dedicated to the personal security of individuals (Atkins 1990).

The total car capacity is the sum of the total seating and the total standing capacity. The total standing capacity is measured as the number of standing individuals per m^2, and provides a good indication of the crowding levels while taking into account crowding expectations per city. Individuals tend to be more fearful in environments where they do not have a clear view of their surroundings (Loukaitou-Sideris and Fink 2009, Zelinka and Brennan 2001): hence we expect that more car capacity increase the level of perceived safety.

Metro ridership defines the number of journeys done by passengers (including fare evaders). The average number of lines by stations and the average metro line length are computed using the 2018 World Metros Figure produced by the International Association of Public Transport (UITP)[4]. An increased metro ridership is expected to have a positive effect on perceived security. It follows Jacobs (1961) which explains that reaching a critical mass of individuals around reduces crime, as there are more "eyes on the street", the area then becomes more self-policing. Finally, more acts of violence increase the individual perceived risk of accident, and is thus expected to decrease the level of perceived safety. Individuals also appear to be fearful of empty train wagons of Transport (1997). To summarise, the metro ridership and the car capacity respectively indicate actual and potential crowding.

The number of violent acts is the sum of two events that are (a) the number of robberies and (b) the number of acts of violence that occur per year in the metro. The impact of the each subcategory on the perception of safety is included in additional regressions in Appendix (Tables 14 and 15), as incivilities are predictive of fear in urban spaces (Rohe and Burby 1988).

The penultimate control is the average number of stations by line. More stations by line implies more time potentially spent in the metro, therefore a longer time when individuals wait in transports. A strand of the literature (Loukaitou-Sideris 1999) finds that longer waits lead to more anxiety and concerns for personal safety. In addition, this anxiety effect is magnified by individuals' tendency to overestimate waiting times (Fan et al. 2016). The number of cars per train is a feature that indicates the potential to exit between two stations in case of emergency;

[4]See more about the data: https://www.uitp.org/world-metro-figures-2018. This dataset covers a time span going from 2013 to 2017 included.

therefore, we expect a positive effect of cars per train on perceived safety.

The specification we use has the same general form as equation 1, but incorporates more independent variables as follows:

$$y_i^* = x_i^1\beta_1 + x_i^2\beta_2 + \sum_{k=1}^{N} \delta_k \; c_k + \sum_{m=2009}^{M} \gamma_m \; t_m + u_i \tag{8}$$

where x_i^1 is a vector of observable individual characteristics as the one used in Table 2 (age, gender, frequency of use, motive of use). The general influence of the metro characteristics is included in the vector x_i^2.

Table 7 presents the ordered probit estimates from the model in equation 8. The table also uses as an outcome the likelihood of recommending the service to someone known to the respondent[5]. The coefficient related to women remains stable to the inclusion of metro characteristics (columns 1 and 2). However, overall satisfaction and the likelihood to recommend the metro to another person encompass a smaller gender difference than security items, which is in line with security being one dimension of overall satisfaction.

We find that more acts of violence, less cars per train, more stations served by line are associated with lower overall customer satisfaction and perceived safety levels. However, the presence of more staff members (both regular staff and police force) has no significant impact on perceived safety. Metro ridership coefficients indicate that individuals feel more secure in metros with more human activity, which is in line with the literature. We note that questions on overall customer satisfaction differ slightly from security questions.

	Stations are secure for me (1)	Trains are secure for me (2)	Overall satisfaction (3)	Likelihood to recommend to another person (4)
Female (Yes=1)	-0.241***	-0.239***	-0.0422***	-0.0156**
	(0.00673)	(0.00673)	(0.00684)	(0.00658)
Total staff in stations	-0.0146	0.0274	-0.118***	-0.0967***
	(0.0211)	(0.0210)	(0.0215)	(0.0210)
Metro ridership	0.00157***	0.00218***	0.00419***	0.00349***
	(0.000214)	(0.000214)	(0.000218)	(0.000209)
Nb of cars per train	0.00726***	0.0141***	0.00593**	-0.000416
	(0.00258)	(0.00257)	(0.00264)	(0.00252)
Avg number stations by line	-0.0453***	-0.0663***	-0.0910***	-0.0751***
	(0.00547)	(0.00546)	(0.00554)	(0.00532)
Total car capacity	0.00450***	0.00484***	0.0128***	0.0125***
	(0.000653)	(0.000652)	(0.000665)	(0.000637)
Nb of violent acts	-0.000205***	-0.000161***	-0.000130***	-0.000136***
	(2.70e-05)	(2.70e-05)	(2.76e-05)	(2.67e-05)
P-value of coeft. female	<0.001	<0.001	<0.001	0.018
Observations	109,630	109,548	109,743	108,057
McFadden Adj.R^2	0.037	0.033	0.076	0.036

Standard errors in parentheses.
*** p<0.01, ** p<0.05, * p<0.1
Source: Customer Satisfaction Surveys, Comet/Nova, own calculations.
Note: All estimates from regressions computed above contain city and year dummies, as well as sociodemographic controls.

Table 7: Effect of metro characteristics on the perception of safety by gender

Results suggest that the presence of more staff in stations is associated with lower overall

[5]The likelihood of recommending the service to someone is on a increasing 10-point scale, where 10 indicates that the individual will recommend the service for sure

satisfaction. This is a counter-intuitive result, perhaps indicative of lower safety satisfaction in less modern / user-friendly systems. We also include a measure of the total number of acts of violence which have a significant negative effects on perception of safety.

6.1 Testing for Gender Heterogeneity

Table 8 presents estimates for heterogeneity in the perception of metro characteristics by gender. Testing for gender heterogeneity is motivated as responses to metro providers' policies differ by subgroup (Yavuz and Welch 2010). To derive results, we run regressions on the subgroup of women.

	Stations are secure for me (1)	Trains are secure for me (2)	Overall satisfaction (3)	Likelihood to recommend to another person (4)
Female (Yes=1)	-	-	-	-
Total staff in stations	0.0160	0.0614*	-0.0855***	-0.0323
	(0.0322)	(0.0322)	(0.0331)	(0.0319)
Metro ridership	0.00167***	0.00251***	0.00680***	0.00469***
	(0.000416)	(0.000416)	(0.000426)	(0.000408)
Nb of cars per train	0.0142***	0.0229***	0.0163***	0.0121***
	(0.00384)	(0.00383)	(0.00395)	(0.00374)
Avg number stations by line	-0.0716***	-0.103***	-0.144***	-0.126***
	(0.00962)	(0.00962)	(0.00977)	(0.00937)
Total car capacity	0.00658***	0.00811***	0.0196***	0.0180***
	(0.00101)	(0.00101)	(0.00103)	(0.000982)
Nb of violent acts	-0.000292***	-0.000261***	-0.000295***	-0.000226***
	(4.43e-05)	(4.43e-05)	(4.55e-05)	(4.38e-05)
Observations	50,866	50,816	50,930	50,155
McFadden Adj.R^2	0.036	0.031	0.084	0.036

Standard errors in parentheses.
*** p<0.01, ** p<0.05, * p<0.1
Source: Customer Satisfaction Surveys, Comet/Nova, own calculations.
Note: All estimates from regressions computed above contain city and year dummies, as well as sociodemographic controls.

Table 8: Heterogeneity by gender in the perception of metro characteristics

Questions regarding safety have similar outcomes: the positive effects on the perception of total car capacity and the negative effect of acts of violence are larger. The coefficient related to the number of acts of violence is in line with findings presented in the economics of crime literature. Crouch (2009) argues that harassment in public transportation constrains women's freedom of movement: our results suggest that this mechanism goes through the channel of perceptions, as more acts of violence decreases women's feeling of safety. However, the number of staff in stations has a significant and positive effect on safety for women, which indicates different preferences and needs for this subgroup. Finally, the effect of metro ridership effect is still observed: more riders make women feel safer. As found previously, results from general customer satisfaction questions differ slightly from the safety ones. The significant elements are again the same as those observed for the whole population. In a second step, Tables 14 and 15 in Appendix provide a detailed specification for acts of violence, respectively for all individuals and for women. It appears that the total number of robberies particularly affects women's perception of safety negatively.

Estimates in Tables 7 and 8 indicate that metro characteristics have a significant impact on individuals as well as a differentiated impact on women. Our results contribute to the growing strand of literature which shows that environmental factors highly influence women's perception of safety (Whitzman 2012). Our findings also corroborate previous results from the literature indicating heterogeneous responses to urban characteristics by gender (Yavuz and Welch 2010, Loukaitou-Sideris and Fink 2009). The magnitude of the coefficients is quite small, which is in line with Koskela and Pain (2000), and indicates that the gap between women and men in perceived safety cannot exactly be bridged only by changing the metro characteristics as it would be likely to come at a high cost. Therefore, results indicate that "designing out" the fear of urban crime will depend essentially on what type of customer is targeted by a given policy. Yavuz and Welch (2010) and Loukaitou-Sideris and Fink (2009) make a case of the existence of gender discrepancies in responses to the evolution of public transport characteristics. Yet, our estimates provide a more precise understanding of these discrepancies as they indicate that the women's responses are always of a larger magnitude compared to men's. Therefore, this indicates that initiatives to make transport and urban spaces safer for women would also in turn make them safer for everyone (Viswanath and Basu 2015).

By capturing a gap in the perception of safety between men and women, our estimates provide evidence that in an inner-city context, it is not just crime (Twinam 2017, Phillips and Sandler 2015) but the perception of crime that decreases with the improvement of public transportation. This suggests that there are ways in which operators can intervene to improve perceptions of safety with gender specific concerns in mind, but also help ruling out some tools highlighted by qualitative analyses (Vanier and d'Arbois de Jubainville 2017). We also note that inclusion of the additional variables does not bridge the gap in stated safety levels between men and women.

Overall, women are not deeply dissatisfied with the levels of safety in public transport. The general distribution of safety responses in Fig. 8 and 9 indicate that women are on average satisfied with metro and bus services. We find that 45% of the female population feel safe in metro stations and trains, while 55% of women feel secure in buses. Therefore, the gap between men and women likely encompasses general differences in the overall perception of transport and the urban environment, aside from the intrinsic network characteristics.

7 Conclusion

This paper quantifies gender differences in the perception of safety and satisfaction in public transpor, using large-scale unique customer satisfaction surveys worldwide. Results indicate a significant gender gap in the perception of safety: women are 10% more likely than men to feel unsafe in metros and 6% more likely to feel unsafe in buses. This gender gap is larger for safety than for overall satisfaction (3% in metros and 2.5% in buses), which is consistent with safety being one important dimension of overall satisfaction. Results are stable and robust across various model specifications and we find that effects are heterogeneous with respect to age and country. Metro characteristics appear to have an influence on perceived safety: more acts of violence, less cars per trains and emptier vehicles decrease the feeling of safety among women; while the presence of more staff in metros does not significantly increase female safety. These results suggest grounds for intervention for service providers (*e.g.*, more staff in stations or information campaigns to reduce the number of violent acts, see Vanier and d'Arbois de Jubainville (2017)).

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 characteristics of transport provision and the evaluation of safety by gender. Metro characteristics have a significant impact on individuals' perceived safety and satisfaction, and the magnitude of those effects is always magnified for women. However, while metro characteristics are important, we also find that they cannot fully explain the gender gap that we observe in perceptions of safety which is likely due to other general differences in gender perceptions of the urban environment.

References

Atkins, S. (1990). Personal security as a transport issue: a state-of-the-art review. *Transport Reviews 10*(2), 111–125.

Carter, M. (2005). Gender differences in experience with and fear of crime in relation to public transport. *Research on women's issues in transportation 2*, 100.

Condon, S., M. Lieber, and F. Maillochon (2007). Feeling Unsafe in Public Places: Understanding Women's Fears.

Crouch, M. (2009, 7). Sexual Harassment in Public Places. *Social Philosophy Today 25*, 137–148.

de Oña, J. and R. de Oña (2014). Quality of service in public transport based on customer satisfaction surveys: A review and assessment of methodological approaches. *Transportation Science 49*(3), 605–622.

Deaton, A. (1985). Panel data from time series of cross-sections. *Journal of econometrics 30*(1-2), 109–126.

Delbosc, A. and G. Currie (2012). Modelling the causes and impacts of personal safety perceptions on public transport ridership. *Transport Policy 24*, 302–309.

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.

European Commission (2014). She moves - Women's Issues in Transportation. Technical report.

Fan, Y., A. Guthrie, and D. Levinson (2016). Waiting time perceptions at transit stops and stations: Effects of basic amenities, gender, and security. *Transportation Research Part A: Policy and Practice 88*, 251–264.

Fellesson, M. and M. Friman (2012). Perceived satisfaction with public transport service in nine european cities. In *Journal of the Transportation Research Forum*, Volume 47.

Fiorio, C. V., M. Florio, and G. Perucca (2011). Consumers' satisfaction and regulation of local public transport: evidence from european cities.

for Transport, U. D. (2004). *People's perceptions of personal security and their concerns about crime on public transport: research findings*. Department for Transport.

Gilchrist, E., J. Bannister, J. Ditton, and S. Farrall (1998). Women and the 'fear of crime' challenging the accepted stereotype. *The British Journal of Criminology 38*(2), 283–298.

Hall, A. R. (2005). *Generalized Method of Moments*. Advanced Texts in Econometrics. Oxford: Oxford University Press.

Hall, R. E. (1985). *Ask Any Woman: A London Inquiry Into Rape and Sexual Assault*. Falling Wall Press; Bristol.

Hernandez, S. and A. Monzon (2016). Key factors for defining an efficient urban transport interchange: Users' perceptions. *Cities 50*, 158–167.

Jacobs, J. (1961). *The death and life of great American cities* (1 ed.). New York: Random Books.

Johnson, M. D., A. Herrmann, and A. Gustafsson (2002). Comparing customer satisfaction across industries and countries. *Journal of Economic Psychology 23*(6), 749–769.

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.

Koskela, H. and R. Pain (2000). Revisiting fear and place: women's fear of attack and the built environment. *Geoforum 31*(2), 269–280.

Loukaitou-Sideris, A. (1999). Hot spots of bus stop crime: The importance of environmental attributes. *Journal of the American Planning association 65*(4), 395–411.

Loukaitou-Sideris, A. and C. Fink (2009). Addressing women's fear of victimization in transportation settings: A survey of us transit agencies. *Urban Affairs Review 44*(4), 554–587.

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

Madge, C. (1997). Public parks and the geography of fear. *Tijdschrift voor economische en sociale geografie 88*(3), 237–250.

Mannering, F. L., V. Shankar, and C. R. Bhat (2016). Unobserved heterogeneity and the statistical analysis of highway accident data. *Analytic methods in accident research 11*, 1–16.

Namgung, M. and G. Akar (2014). Role of gender and attitudes on public transportation use. *Transportation research record 2415*(1), 136–144.

O'Brien, E. (2005). Publics* and woodlands in england: well-being, local identity, social learning, conflict and management. *Forestry 78*(4), 321–336.

of Transport, D. (1997). *Crime concern and transport and travel research: Perceptions of safety from crime on public transpor*. London:Author.

Oliver, R. L. (1997). *Satisfaction: A behavioral perspective on the consumer*. Irwin/McGraw-Hill, New York.

Pain, R. (2001). Gender, race, age and fear in the city. *Urban studies 38*(5-6), 899–913.

Pain, R. H. (1997). Social geographies of women's fear of crime. *Transactions of the Institute of British geographers*, 231–244.

Phillips, D. C. and D. Sandler (2015, 5). Does public transit spread crime? Evidence from temporary rail station closures. *Regional Science and Urban Economics 52*, 13–26.

Riger, S. and M. Gordon (1989). The female fear.

Rohe, W. M. and R. J. Burby (1988). Fear of crime in public housing. *Environment and Behavior 20*(6), 700–720.

Ross, A. (2010, 1). Violent crime, entrepreneurship, and cities. *Journal of Urban Economics 67*(1), 135–149.

Schulz, D. and S. Gilbert (1996). Women and transit security: A new look at an old issue. In *Women's Travel Issues Second National Conference*, pp. 550–562. Citeseer.

Shaw, M. (2002). Gender and crime prevention. Technical report, International Observer.

St-Louis, E., K. Manaugh, D. van Lierop, and A. El-Geneidy (2014). The happy commuter: a comparison of commuter satisfaction across modes. *Transportation research part F: traffic psychology and behaviour 26*, 160–170.

Stanko, E. (1990). Everyday violence: Women's and men's experience of personal danger. *London: Pandora*.

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.

Trompet, M., R. J. Anderson, and D. J. Graham (2018). Improved understanding of the relative quality of bus public transit using a balanced approach to performance data normalization. *Transportation Research Part A: Policy and Practice 114*, 13–23.

Trompet, M., R. Parasram, and R. J. Anderson (2013). Benchmarking disaggregate customer satisfaction scores of bus operators in different cities and countries. *Transportation Research Record 2351*(1), 14–22.

Twinam, T. (2017, 7). Danger zone: Land use and the geography of neighborhood crime. *Journal of Urban Economics 100*, 104–119.

Tyrinopoulos, Y. and C. Antoniou (2008). Public transit user satisfaction: Variability and policy implications. *Transport Policy 15*(4), 260–272.

Ulfarsson, G. F. and F. L. Mannering (2004). Differences in male and female injury severities in sport-utility vehicle, minivan, pickup and passenger car accidents. *Accident Analysis & Prevention 36*(2), 135–147.

UN Women (2016). Gender Statistics - Violence against Women.

Valentine, G. (1990). Women's fear and the design of public space. *Built Environment (1978-)*, 288–303.

Vanier, C. and H. d'Arbois de Jubainville (2017, 9). Feeling unsafe in public transportation: A profile analysis of female users in the Parisian region. *Crime Prevention and Community Safety 19*(3-4), 251–263.

Viswanath, K. and A. Basu (2015, 1). SafetiPin: an innovative mobile app to collect data on women's safety in Indian cities. *Gender & Development 23*(1), 45–60.

Washington, S. P., M. G. Karlaftis, and F. Mannering (2010). *Statistical and econometric methods for transportation data analysis.* Chapman and Hall/CRC.

Wekerle, G. R. and C. Whitzman (1995). *Safe cities: Guidelines for planning, design, and management.* Van Nostrand Reinhold New York.

Whitzman, C. (2012, 5). *The Handbook of Community Safety Gender and Violence Prevention.* Routledge.

Yavuz, N. and E. W. Welch (2010). Addressing fear of crime in public space: Gender differences in reaction to safety measures in train transit. *Urban studies 47*(12), 2491–2515.

Zelinka, A. and D. Brennan (2001). *SafeScape. Creating Safer, More Livable Communities through Planning and Design.*

A Appendix A: CoMET, Nova and IBBG groups

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 IBBG group is a large consortium of major bus networks including: Transport Metropolitans de Barcelona (TMB, Barcelona), Société des Transports Intercommunaux de Bruxelles (STIB, Brussels), Dublin Bus (Dublin), IETT Isletmeleri Genel Müdürlügü (IETT, Istanbul), Rapid Bus Sdn Bhd (Rapid KL, Kuala Lumpur), Companhia Carris de Ferro de Lisboa (Carris, Lisbon), London Buses (LBSL, London), Société de Transport de Montréal (STM, Montréal), MTA New York City Transit and MTA Bus (New York), the Régie Autonome des Transports Parisiens (RATP, Paris), King County Metro (KCM, Seattle), SMRT Buses (Singapore), State Transit (Sydney), and Coast Mountain Bus Company (CMBC, Vancouver).

B Appendix B: 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 9: Descriptive Statistics - Metro Data

	Transport Mode: Bus										
	Years										
	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018	2009/18
Women	0.50	0.47	0.54	0.54	0.52	0.53	0.53	0.53	0.51	0.51	0.52
	(0.50)	(0.50)	(0.50)	(0.50)	(0.50)	(0.50)	(0.50)	(0.50)	(0.50)	(0.50)	(0.50)
Age											
Less than 18	0.04	0.06	0.05	0.04	0.03	0.04	0.03	0.05	0.05	0.06	0.05
	(0.19)	(0.24)	(0.22)	(0.20)	(0.18)	(0.19)	(0.18)	(0.22)	(0.22)	(0.24)	(0.21)
18-29	0.37	0.38	0.36	0.36	0.32	0.32	0.30	0.35	0.34	0.31	0.34
	(0.48)	(0.49)	(0.48)	(0.48)	(0.47)	(0.47)	(0.46)	(0.48)	(0.47)	(0.46)	(0.47)
30-39	0.23	0.25	0.21	0.23	0.22	0.22	0.23	0.23	0.24	0.21	0.22
	(0.42)	(0.43)	(0.41)	(0.42)	(0.41)	(0.42)	(0.42)	(0.42)	(0.42)	(0.41)	(0.42)
40-49	0.17	0.17	0.16	0.15	0.16	0.15	0.17	0.15	0.16	0.17	0.16
	(0.37)	(0.37)	(0.36)	(0.36)	(0.37)	(0.36)	(0.37)	(0.36)	(0.36)	(0.37)	(0.37)
50-65	0.19	0.14	0.18	0.18	0.21	0.20	0.21	0.17	0.17	0.19	0.19
	(0.39)	(0.35)	(0.38)	(0.38)	(0.41)	(0.40)	(0.41)	(0.37)	(0.37)	(0.39)	(0.39)
More than 65	0.00	0.00	0.04	0.04	0.06	0.06	0.05	0.05	0.05	0.06	0.05
	(0.02)	(0.00)	(0.20)	(0.19)	(0.23)	(0.24)	(0.22)	(0.22)	(0.21)	(0.24)	(0.22)
Frequency of use											
Very Often	0.55	0.52	0.59	0.57	0.52	0.55	0.59	0.60	0.58	0.56	0.57
	(0.50)	(0.50)	(0.49)	(0.50)	(0.50)	(0.50)	(0.49)	(0.49)	(0.49)	(0.50)	(0.50)
Often (min: 3 times/week)	0.20	0.19	0.20	0.22	0.23	0.24	0.23	0.24	0.24	0.25	0.23
	(0.40)	(0.40)	(0.40)	(0.41)	(0.42)	(0.43)	(0.42)	(0.43)	(0.43)	(0.43)	(0.42)
Sometimes (min: once/week)	0.13	0.12	0.13	0.13	0.15	0.13	0.11	0.11	0.13	0.13	0.13
	(0.34)	(0.33)	(0.33)	(0.33)	(0.35)	(0.34)	(0.32)	(0.31)	(0.33)	(0.34)	(0.33)
Rarely (min: once/month)	0.07	0.07	0.05	0.06	0.07	0.05	0.04	0.04	0.04	0.04	0.05
	(0.25)	(0.26)	(0.21)	(0.23)	(0.25)	(0.22)	(0.20)	(0.19)	(0.19)	(0.20)	(0.21)
Very rarely (less than once/month)	0.06	0.09	0.03	0.03	0.03	0.03	0.02	0.02	0.02	0.02	0.03
	(0.23)	(0.28)	(0.17)	(0.17)	(0.18)	(0.17)	(0.14)	(0.14)	(0.13)	(0.14)	(0.16)
Most frequent travel motive											
Work/Education	0.74	0.69	0.75	0.73	0.70	0.72	0.79	0.77	0.75	0.73	0.74
	(0.44)	(0.46)	(0.43)	(0.44)	(0.46)	(0.45)	(0.41)	(0.42)	(0.43)	(0.44)	(0.44)
Shopping	0.04	0.03	0.05	0.05	0.05	0.05	0.03	0.04	0.04	0.04	0.04
	(0.20)	(0.17)	(0.21)	(0.21)	(0.21)	(0.21)	(0.18)	(0.20)	(0.21)	(0.20)	(0.20)
Leisure	0.16	0.16	0.15	0.18	0.20	0.17	0.13	0.14	0.15	0.17	0.16
	(0.37)	(0.37)	(0.36)	(0.38)	(0.40)	(0.38)	(0.34)	(0.34)	(0.36)	(0.38)	(0.37)
Doctor	0.02	0.03	0.02	0.02	0.02	0.03	0.02	0.03	0.03	0.03	0.03
	(0.15)	(0.16)	(0.14)	(0.13)	(0.15)	(0.17)	(0.15)	(0.16)	(0.17)	(0.17)	(0.16)
Other	0.03	0.10	0.03	0.03	0.02	0.03	0.03	0.02	0.03	0.02	0.03
	(0.17)	(0.30)	(0.17)	(0.16)	(0.15)	(0.18)	(0.16)	(0.15)	(0.17)	(0.16)	(0.17)

Standard errors in parentheses.
Source: IBBG data.

Table 10: Descriptive Statistics - Bus Data

C Appendix C: Distribution of Probabilities

	Stations are secure for me (1)	Trains are secure for me (2)	Overall Satisfaction (4)	Likelihood to recommend to another person (3
Share of female respondents	**-0.409***	**-0.661*****	**-0.134**	**-0.179**
	(0.218)	**(0.239)**	**(0.213)**	**(0.185)**
Cutpoint 1	-2.417***	-2.553***	-2.001***	-2.104***
	(0.254)	(0.275)	(0.358)	(0.352)
Cutpoint 2	-1.624***	-1.810***	-0.731**	-1.941***
	(0.254)	(0.241)	(0.361)	(0.334)
Cutpoint 3	0.703***	0.540**	1.100***	-1.797***
	(0.264)	(0.240)	(0.293)	(0.298)
Cutpoint 4	2.904***	2.752***	3.437***	-1.401***
	(0.329)	(0.275)	(0.287)	(0.260)
Cutpoint 5				-0.919***
				(0.257)
Cutpoint 6				-0.222
				(0.272)
Cutpoint 7				0.574**
				(0.284)
Cutpoint 8				1.622***
				(0.308)
Cutpoint 9				2.661***
				(0.283)
Cutpoint 10				3.464***
				(0.280)
Var (city level)	0.423***	0.388***	0.467***	0.391**
	(0.105)	(0.0848)	(0.0933)	(0.157)
Var (female) at city level	0.423***	0.388***	0.467***	0.453**
	(0.105)	(0.0848)	(0.0933)	(0.194)
Observations	2,899	2,899	2,898	2,898
Number of groups	25	25	25	25
Prob chi2(2)	0	0	0	0

Standard errors in parentheses.
*** p<0.01, ** p<0.05, * p<0.1
Source: Customer Satisfaction Surveys, Comet/Nova, own calculations.

Table 11: Gender Coefficients on the Perception of Safety - Multilevel mixed-effects ordered probit estimates (Random Intercept and Random Coefficient)

Dimension	Availability/Convenience			Ease	Information			Reliability	Comfort		Customer Care		
Variable	Freq.	Hours	Network	Interchange	Disruption	During Journey	Plan Journey		Station	Train	Pay Ticket	Answer Questions	Staff Train
	(A1)	(A2)	(A3)	(B1)	(C1)	(C2)	(C3)	(D1)	(E1)	(E2)	(F1)	(F2)	(F3)
Female (Yes=1)	**-0.0428*****	**-0.00270**	**0.0326*****	**-0.0560*****	**-0.0243*****	**-0.0148*****	**-0.0444*****	**-0.0986*****	**-0.0898*****	**-0.0943*****	**-0.0421*****	**-0.0562*****	**-0.0242*****
	(0.00545)	**(0.00547)**	**(0.00543)**	**(0.00541)**	**(0.00548)**	**(0.00539)**	**(0.00556)**	**(0.00546)**	**(0.00541)**	**(0.00539)**	**(0.00549)**	**(0.00570)**	**(0.00550)**
Age (ref: 18-29)													
Less than 18	0.200***	0.495***	0.238***	0.232***	0.330***	0.315***	0.254***	0.249***	0.213***	0.230***	0.211***	0.292***	0.234***
	(0.0146)	(0.0150)	(0.0146)	(0.0146)	(0.0146)	(0.0146)	(0.0153)	(0.0145)	(0.0145)	(0.0144)	(0.0149)	(0.0153)	(0.0149)
30-39	-0.0457***	0.0453***	-0.0638***	-0.113***	-0.0890***	-0.0719***	-0.149***	0.00409	0.00444	0.00399	-0.0697***	-0.0953***	-0.0649***
	(0.00711)	(0.00713)	(0.00708)	(0.00706)	(0.00718)	(0.00704)	(0.00726)	(0.00713)	(0.00708)	(0.00705)	(0.00716)	(0.00744)	(0.00718)
40-49	0.0273***	0.153***	0.0287***	-0.0798***	-0.0326***	-0.0248***	-0.158***	0.110***	0.0730***	0.0742***	-0.0639***	-0.0775***	-0.0333***
	(0.00836)	(0.00841)	(0.00834)	(0.00830)	(0.00840)	(0.00826)	(0.00852)	(0.00837)	(0.00831)	(0.00828)	(0.00842)	(0.00869)	(0.00841)
50-65	0.113***	0.230***	0.119***	-0.0246***	0.0132	0.0167*	-0.140***	0.205***	0.160***	0.118***	0.0304***	-0.0186**	0.0328***
	(0.00868)	(0.00876)	(0.00868)	(0.00863)	(0.00868)	(0.00858)	(0.00886)	(0.00868)	(0.00862)	(0.00859)	(0.00878)	(0.00903)	(0.00873)
Over 65	0.284***	0.452***	0.267***	0.0735***	0.0928***	0.107***	-0.0690***	0.356***	0.355***	0.289***	0.133***	0.121***	0.198***
	(0.0161)	(0.0164)	(0.0161)	(0.0159)	(0.0159)	(0.0158)	(0.0164)	(0.0160)	(0.0158)	(0.0157)	(0.0164)	(0.0166)	(0.0161)
Main travel purpose (ref: Work)													
Education	0.0690***	-0.0677***	0.00820	-0.00925	0.101***	0.0762***	-0.000490	0.114***	0.0120	0.0599***	-0.0556***	0.0641***	0.0398***
	(0.0117)	(0.0118)	(0.0117)	(0.0116)	(0.0117)	(0.0116)	(0.0119)	(0.0117)	(0.0116)	(0.0116)	(0.0118)	(0.0121)	(0.0118)
Shopping	0.0959***	0.0406***	0.0676***	0.114***	0.127***	0.0921***	0.142***	0.128***	0.0819***	0.0825***	0.0663***	0.131***	0.0915***
	(0.00864)	(0.00866)	(0.00860)	(0.00860)	(0.00871)	(0.00856)	(0.00889)	(0.00866)	(0.00857)	(0.00854)	(0.00873)	(0.00907)	(0.00878)
Leisure	0.178***	0.00825	0.150***	0.143***	0.220***	0.175***	0.187***	0.236***	0.107***	0.136***	0.101***	0.164***	0.135***
	(0.0101)	(0.0101)	(0.0100)	(0.0100)	(0.0100)	(0.00997)	(0.0103)	(0.0101)	(0.00997)	(0.00993)	(0.0102)	(0.0106)	(0.0102)
Doctor	0.127***	0.00680	0.0927***	0.0552***	0.195***	0.147***	0.114***	0.168***	0.0628***	0.0989***	0.0315**	0.113***	0.0846***
	(0.0130)	(0.0131)	(0.0130)	(0.0129)	(0.0130)	(0.0129)	(0.0133)	(0.0130)	(0.0129)	(0.0129)	(0.0131)	(0.0134)	(0.0131)
Other	0.0327	-0.0996***	0.0551***	-0.0196	0.131***	0.103***	0.0473**	0.102***	0.0234	0.0522**	-0.0104	0.0502**	0.0475**
	(0.0213)	(0.0215)	(0.0213)	(0.0211)	(0.0214)	(0.0212)	(0.0219)	(0.0213)	(0.0212)	(0.0211)	(0.0216)	(0.0223)	(0.0215)
Frequency use trains (ref: very often)													
Often	0.148***	0.0658***	-0.0512***	0.0104	0.110***	0.0860***	0.0251***	0.167***	0.0822***	0.124***	0.00285	0.0767***	0.0682***
	(0.00741)	(0.00746)	(0.00740)	(0.00738)	(0.00743)	(0.00734)	(0.00759)	(0.00742)	(0.00736)	(0.00733)	(0.00750)	(0.00775)	(0.00749)
Sometimes	0.205***	0.103***	-0.102***	0.00983	0.140***	0.0980***	0.00598	0.213***	0.112***	0.168***	-0.0269**	0.0728***	0.0530***
	(0.0105)	(0.0106)	(0.0105)	(0.0105)	(0.0105)	(0.0104)	(0.0108)	(0.0105)	(0.0104)	(0.0104)	(0.0106)	(0.0110)	(0.0106)
Rarely	0.273***	0.184***	-0.120***	0.0220	0.206***	0.153***	0.0249*	0.269***	0.162***	0.216***	-0.0189	0.138***	0.115***
	(0.0136)	(0.0137)	(0.0135)	(0.0135)	(0.0136)	(0.0135)	(0.0139)	(0.0136)	(0.0135)	(0.0134)	(0.0137)	(0.0144)	(0.0138)
Very Rarely	0.319***	0.176***	-0.143***	0.0305*	0.296***	0.218***	-0.0989***	0.255***	0.180***	0.234***	-0.0780***	0.184***	0.125***
	(0.0174)	(0.0175)	(0.0172)	(0.0172)	(0.0176)	(0.0174)	(0.0176)	(0.0174)	(0.0172)	(0.0171)	(0.0174)	(0.0184)	(0.0177)
P-value of coeft. female	<0.001	0.622	<0.001	<0.001	<0.001	0.006	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001
Observations	169,813	169,697	169,667	169,385	166,097	167,753	169,431	169,728	169,931	169,951	168,984	150,099	161,733
McFadden Adj.R^2	0.087	0.049	0.029	0.023	0.055	0.040	0.030	0.090	0.094	0.074	0.027	0.044	0.033

Standard errors in parentheses.
*** p<0.01, ** p<0.05, * p<0.1
Note: All estimates from regressions computed above contain city and year dummies.

Table 12: Ordered Probit Estimations: Other Dimensions for Metros

Dimension	Availability	Ease		Information		Reliability	Comfort				Customer care		
Variable	Network	Interchange	Move in bus	Time	Services		Good ride	Spacious	Good Light	Well Dressed	Helpful	Pay Ticket	Sort Pbs
	(A1)	(B1)	(B2)	(C1)	(C2)	(D1)	(E1)	(E2)	(E3)	(E4)	(F1)	(F2)	(F3)
Female (Yes=1)	**-0.0429*****	**-0.159*****	**-0.124*****	**-0.0194*****	**-0.0499*****	**-0.0554*****	**-0.103*****	**0.0128****	**-0.132*****	**0.0879*****	**-0.0254*****	**-0.0380*****	**-0.0335*****
	(0.00515)	**(0.00526)**	**(0.00509)**	**(0.00509)**	**(0.00513)**	**(0.00515)**	**(0.00508)**	**(0.00511)**	**(0.00506)**	**(0.00533)**	**(0.00513)**	**(0.00523)**	**(0.00538)**
Age (ref: 18-29)													
Less than 18	0.287***	0.0936***	0.119***	0.315***	0.227***	0.300***	0.360***	0.173***	0.198***	0.331***	0.217***	0.190***	0.363***
	(0.0127)	(0.0129)	(0.0124)	(0.0125)	(0.0127)	(0.0126)	(0.0125)	(0.0125)	(0.0124)	(0.0135)	(0.0127)	(0.0129)	(0.0131)
30-39	-0.0646***	-0.0923***	-0.0732***	-0.00869	-0.0355***	0.0109	-0.0812***	-0.0413***	-0.0231***	-0.198***	-0.0611***	0.00265	-0.101***
	(0.00691)	(0.00705)	(0.00685)	(0.00685)	(0.00688)	(0.00690)	(0.00682)	(0.00688)	(0.00681)	(0.00719)	(0.00690)	(0.00700)	(0.00723)
40-49	-0.0119	-0.0613***	-0.0434***	0.0865***	0.0153**	0.145***	-0.0528***	-0.0629***	-0.0126	-0.289***	-0.0416***	0.106***	-0.109***
	(0.00784)	(0.00801)	(0.00773)	(0.00775)	(0.00781)	(0.00784)	(0.00772)	(0.00776)	(0.00770)	(0.00812)	(0.00781)	(0.00794)	(0.00818)
50-65	0.0485***	-0.0919***	-0.0873***	0.126***	0.0114	0.251***	-0.0568***	-0.106***	-0.0316***	-0.388***	-0.0345***	0.206***	-0.121***
	(0.00771)	(0.00788)	(0.00759)	(0.00761)	(0.00766)	(0.00773)	(0.00758)	(0.00761)	(0.00755)	(0.00798)	(0.00767)	(0.00784)	(0.00804)
Over 65	0.168***	-0.0959***	-0.0621***	0.175***	0.0479***	0.377***	0.0585***	0.0663***	0.152***	-0.378***	0.150***	0.379***	-0.00356
	(0.0133)	(0.0135)	(0.0130)	(0.0131)	(0.0132)	(0.0133)	(0.0130)	(0.0131)	(0.0129)	(0.0136)	(0.0132)	(0.0136)	(0.0139)
Main travel purpose (ref: work/education)													
Shopping	0.149***	-0.0369***	0.0413***	0.166***	0.0727***	0.158***	0.0612***	0.125***	0.131***	0.0584***	0.0684***	0.0329**	0.137***
	(0.0134)	(0.0136)	(0.0131)	(0.0132)	(0.0133)	(0.0134)	(0.0132)	(0.0132)	(0.0131)	(0.0138)	(0.0133)	(0.0136)	(0.0140)
Leisure	0.135***	0.103***	0.118***	0.137***	0.111***	0.169***	0.0985***	0.142***	0.127***	0.0712***	0.104***	0.0818***	0.118***
	(0.00833)	(0.00856)	(0.00820)	(0.00822)	(0.00829)	(0.00834)	(0.00820)	(0.00822)	(0.00815)	(0.00861)	(0.00830)	(0.00844)	(0.00881)
Doctor	0.0331*	-0.248***	-0.130***	0.0566***	-0.0345**	0.0464***	-0.0365**	-0.0231	0.0537***	0.0108	-0.0439***	-0.0750***	-0.0152
	(0.0170)	(0.0170)	(0.0167)	(0.0168)	(0.0169)	(0.0170)	(0.0167)	(0.0168)	(0.0167)	(0.0176)	(0.0169)	(0.0172)	(0.0176)
Other	-0.0113	-0.112***	-0.0840***	0.00534	-0.0506***	0.0300*	-0.0384**	-0.00137	0.0389**	0.0195	0.00727	-0.0553***	0.00684
	(0.0157)	(0.0160)	(0.0155)	(0.0156)	(0.0156)	(0.0158)	(0.0155)	(0.0155)	(0.0154)	(0.0163)	(0.0157)	(0.0159)	(0.0165)
Frequency of use (ref: very often)													
Often	-0.0412***	0.119***	0.102***	0.0586***	0.0313***	0.103***	0.0980***	0.119***	0.127***	0.0276***	0.0744***	-0.0615***	0.113***
	(0.00653)	(0.00667)	(0.00644)	(0.00645)	(0.00650)	(0.00653)	(0.00643)	(0.00647)	(0.00641)	(0.00676)	(0.00651)	(0.00663)	(0.00682)
Sometimes	-0.132***	0.184***	0.190***	0.0635***	0.0310***	0.132***	0.179***	0.179***	0.200***	0.0412***	0.112***	-0.106***	0.153***
	(0.00878)	(0.00903)	(0.00867)	(0.00869)	(0.00875)	(0.00881)	(0.00867)	(0.00870)	(0.00862)	(0.00911)	(0.00878)	(0.00891)	(0.00933)
Rarely	-0.308***	0.186***	0.223***	0.0330**	-0.00907	0.102***	0.232***	0.184***	0.215***	0.0177	0.0883***	-0.179***	0.150***
	(0.0129)	(0.0134)	(0.0128)	(0.0128)	(0.0129)	(0.0130)	(0.0128)	(0.0128)	(0.0127)	(0.0134)	(0.0130)	(0.0131)	(0.0141)
Very rarely	-0.581***	0.0363**	0.208***	0.0150	-0.143***	0.00812	0.207***	0.154***	0.214***	-0.0506***	-0.00438	-0.306***	0.0735***
	(0.0169)	(0.0173)	(0.0167)	(0.0169)	(0.0168)	(0.0171)	(0.0167)	(0.0167)	(0.0166)	(0.0174)	(0.0169)	(0.0170)	(0.0183)
P-value of coeft. female	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001	0.012	<0.001	<0.001	<0.001	<0.001	<0.001
Observations	180,003	180,339	180,302	177,514	179,387	180,073	180,236	180,263	180,208	176,802	178,089	177,149	159,596
McFadden Adj.R^2	0.014	0.043	0.051	0.019	0.021	0.029	0.020	0.043	0.013	0.054	0.022	0.026	0.025

Standard errors in parentheses.
*** p<0.01, ** p<0.05, * p<0.1
Note: All estimates from regressions computed above contain city and year dummies.

Table 13: Ordered Probit Estimations: Other Dimensions for Buses

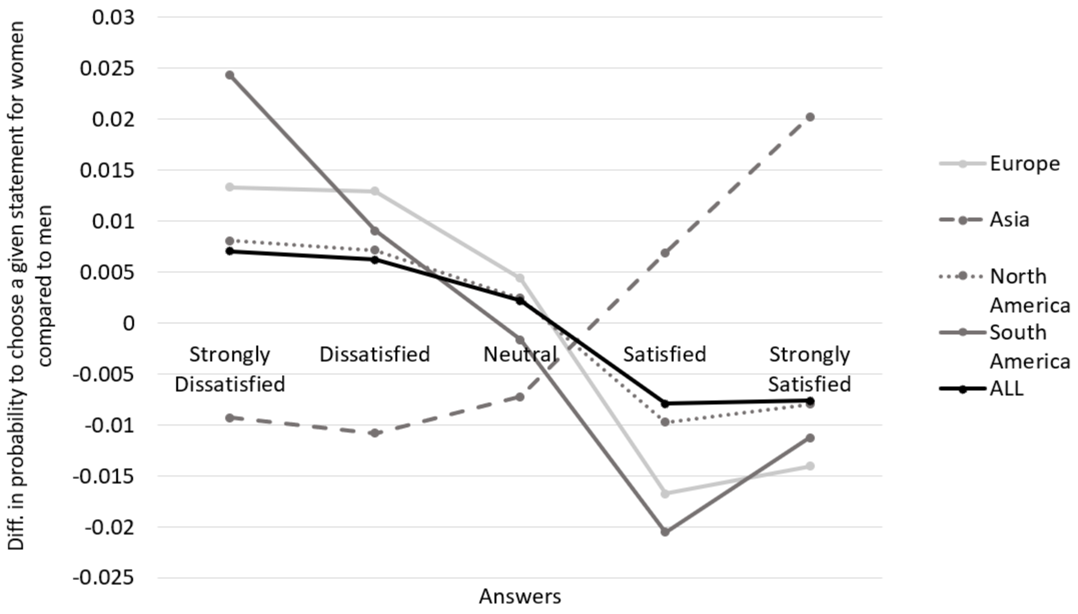


Figure 6: Regional Marginal Effects for the Overall Satisfaction with the Metro System

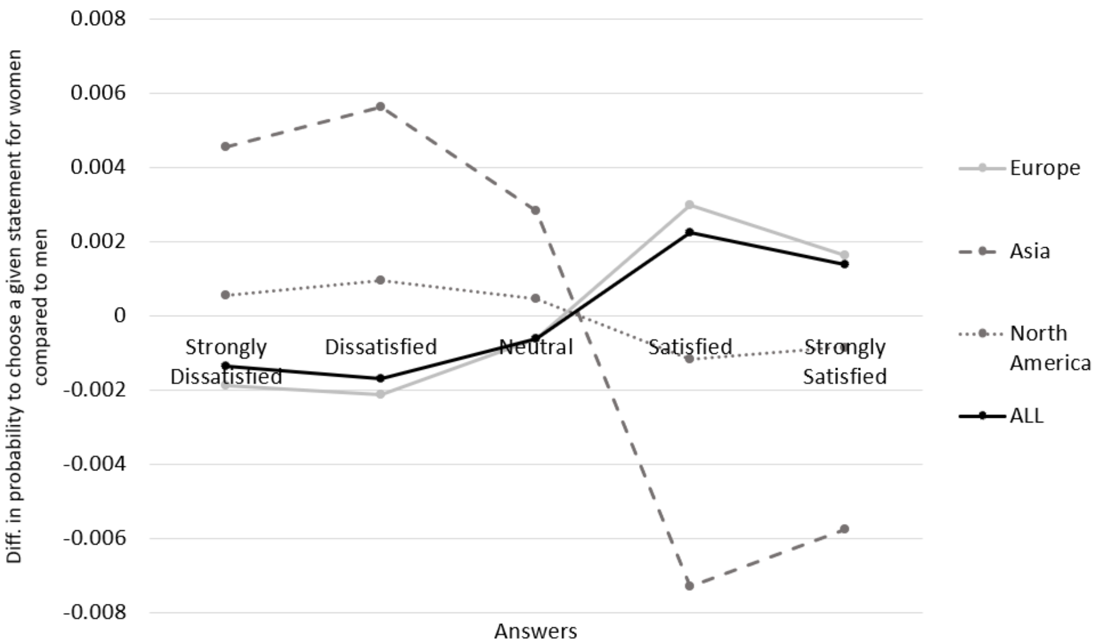


Figure 7: Regional Marginal Effects for the Overall Satisfaction with the Bus System

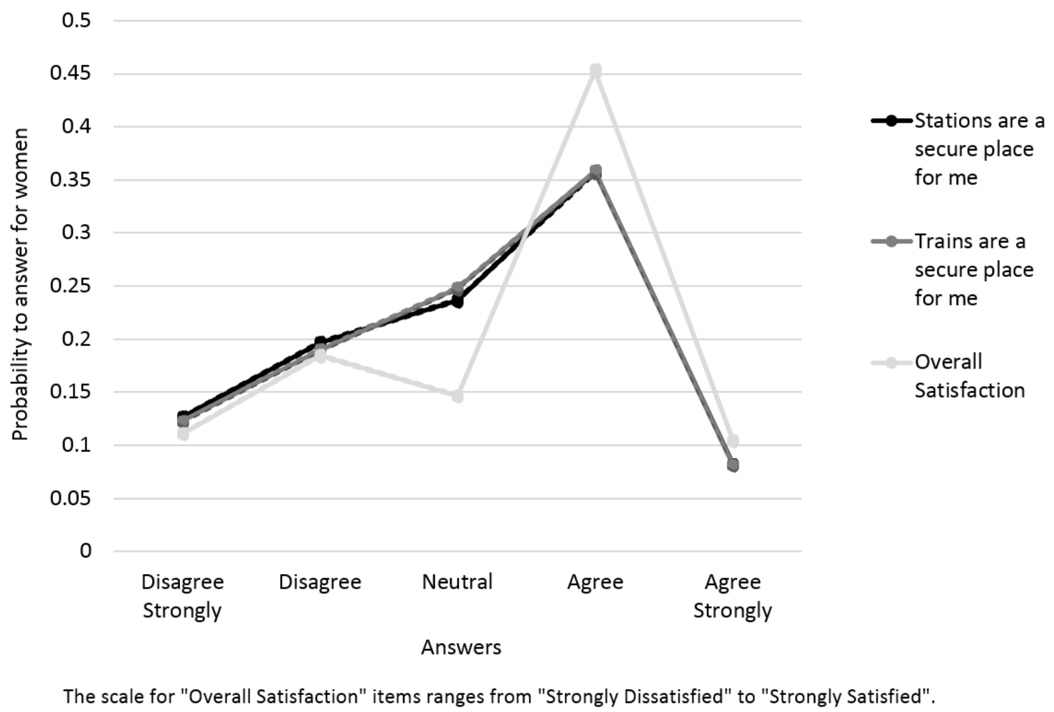


Figure 8: General distribution of safety responses for women - Metro Data

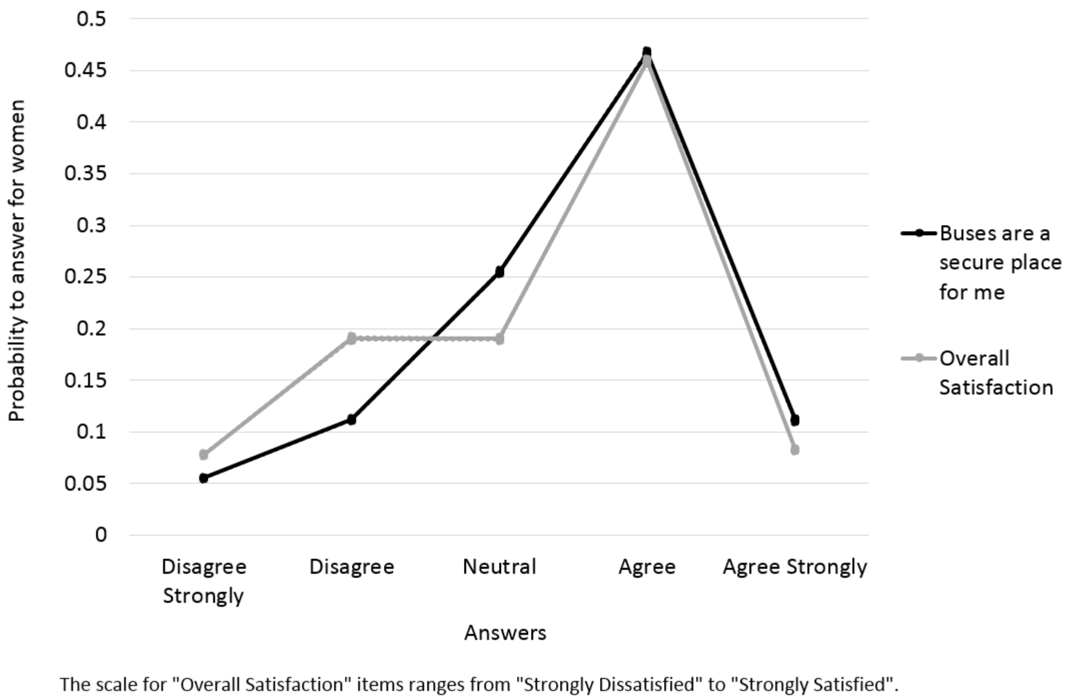


Figure 9: General distribution of safety responses for women - Bus Data

D Appendix D: Influence of Metro Characteristics

	Stations are secure for me (1)	Trains are secure for me (2)	Overall satisfaction (3)	Likelihood to recommend to another person (4)
Female (Yes=1)	**-0.241*** **	**-0.239*** **	**-0.0431*** **	**-0.0162** **
	(0.00673)	**(0.00673)**	**(0.00684)**	**(0.00658)**
Total staff in stations	-0.0256	0.0188	-0.185***	-0.143***
	(0.0221)	(0.0221)	(0.0226)	(0.0221)
Metro ridership	0.00157***	0.00218***	0.00417***	0.00348***
	(0.000214)	(0.000214)	(0.000218)	(0.000209)
Nb of cars per train	0.00742***	0.0142***	0.00692***	0.000283
	(0.00258)	(0.00258)	(0.00264)	(0.00252)
Avg number stations by line	-0.0443***	-0.0654***	-0.0846***	-0.0708***
	(0.00551)	(0.00550)	(0.00558)	(0.00536)
Total car capacity	0.00449***	0.00484***	0.0128***	0.0125***
	(0.000653)	(0.000652)	(0.000665)	(0.000637)
Nb of acts of violence	-0.000230***	-0.000181***	-0.000285***	-0.000243***
	(3.12e-05)	(3.12e-05)	(3.19e-05)	(3.10e-05)
Nb robberies	-0.000111*	-8.77e-05	0.000441***	0.000244***
	(6.42e-05)	(6.42e-05)	(6.53e-05)	(6.20e-05)
P-value of coeft. female	<0.001	<0.001	<0.001	0.014
Observations	109,630	109,548	109,743	108,057
McFadden Adj.R^2	0.037	0.033	0.077	0.036

Standard errors in parentheses.
*** p<0.01, ** p<0.05, * p<0.1
Source: Customer Satisfaction Surveys, Comet/Nova, own calculations.
Note: All estimates from regressions computed above contain city and year dummies, as well as sociodemographic controls.

Table 14: Heterogeneity by gender in the perception of metro characteristics - Detailed Specification for Acts of violence

	Stations are secure for me (1)	Trains are secure for me (2)	Overall satisfaction (3)	Likelihood to recommend to another person (4)
Female (Yes=1)	-	-	-	-
Total staff in stations	0.0158	0.0641*	-0.169***	-0.0873***
	(0.0337)	(0.0337)	(0.0346)	(0.0334)
Metro ridership	0.00167***	0.00250***	0.00705***	0.00487***
	(0.000417)	(0.000417)	(0.000427)	(0.000409)
Nb of cars per train	0.0142***	0.0228***	0.0180***	0.0132***
	(0.00384)	(0.00383)	(0.00396)	(0.00375)
Avg number stations by line	-0.0716***	-0.103***	-0.133***	-0.119***
	(0.00970)	(0.00970)	(0.00986)	(0.00945)
Total car capacity	0.00658***	0.00812***	0.0193***	0.0178***
	(0.00101)	(0.00101)	(0.00103)	(0.000983)
Nb of acts of violence	-0.000293***	-0.000253***	-0.000535***	-0.000384***
	(5.25e-05)	(5.26e-05)	(5.40e-05)	(5.22e-05)
Nb robberies	-0.000291***	-0.000284***	0.000442***	0.000241**
	(9.82e-05)	(9.82e-05)	(0.000100)	(9.50e-05)
Observations	50,866	50,816	50,930	50,155
McFadden Adj.R^2	0.036	0.031	0.085	0.036

Standard errors in parentheses.
*** p<0.01, ** p<0.05, * p<0.1
Source: Customer Satisfaction Surveys, Comet/Nova, own calculations.
Note: All estimates from regressions computed above contain city and year dummies, as well as sociodemographic controls.

Table 15: Heterogeneity by gender in the perception of metro characteristics - Detailed Specification for Acts of violence

	Stations are secure for me (1)	Trains are secure for me (2)	Overall Satisfaction (3)	Likelihood to recommend to another person (4)
Share of female respondents	**-0.226***** **(0.0566)**	**-0.430***** **(0.0685)**	**-0.00905** **(0.0456)**	**-0.128***** **(0.0405)**
Avg. nb. staff in stations	-0.118*** (0.0237)	-0.231*** (0.0260)	-0.414*** (0.0320)	-0.190*** (0.0292)
Avg. metro ridership	0.0106*** (0.00220)	0.0264*** (0.00261)	0.0600*** (0.00350)	0.0480*** (0.00304)
Avg. stations by line	-0.0841*** (0.00764)	-0.0889*** (0.00740)	-0.216*** (0.00799)	-0.212*** (0.00722)
Avg. total capacity	0.00533*** (0.000573)	0.00464*** (0.000635)	0.0114*** (0.00120)	0.0122*** (0.00113)
Avg. number violent acts	-0.000124*** (4.42e-05)	-9.20e-05* (4.71e-05)	0.000711*** (5.81e-05)	0.000522*** (5.34e-05)
Avg. metro ridership	-0.00221*** (0.000766)	-0.00398*** (0.000792)	-0.0196*** (0.00141)	-0.0224*** (0.00132)
Cutpoint 1	-3.289*** (0.178)	-3.553*** (0.183)	-4.880*** (0.250)	-4.944*** (0.228)
Cutpoint 2	-2.692*** (0.177)	-2.927*** (0.183)	-3.640*** (0.249)	-4.808*** (0.227)
Cutpoint 3	-0.364** (0.176)	-0.520*** (0.182)	-1.655*** (0.248)	-4.660*** (0.227)
Cutpoint 4	1.824*** (0.176)	1.808*** (0.182)	0.630** (0.250)	-4.372*** (0.226)
Cutpoint 5				-3.984*** (0.226)
Cutpoint 6				-3.214*** (0.226)
Cutpoint 7				-2.366*** (0.226)
Cutpoint 8				-1.252*** (0.226)
Cutpoint 9				-0.240 (0.226)
Cutpoint 10				0.443* (0.226)
Var (city level)	0.544*** (0.0699)	0.989*** (0.110)	0.203*** (0.0405)	0.164*** (0.0377)
Var (female) at city level	0.298*** (0.0375)	0.379*** (0.0464)	1.686*** (0.202)	1.687*** (0.201)
Observations	1,748	1,748	1,747	1,747
Number of groups	20	20	20	20
Prob chi2(2)	0	0	0	0

Standard errors in parentheses.
*** p<0.01, ** p<0.05, * p<0.1
Source: Customer Satisfaction Surveys, Comet/Nova, own calculations.

Table 16: Effect of Metro Characteristics on the Perception of Safety by Gender - Multilevel Mixed-effects Ordered Probit Estimates (Random Intercept and Random Coefficient)